

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:52:06 ; Search time 94 Seconds
(without alignments)
12132.267 Million cell updates/sec

Title: US-10-051-902A-21

Perfect score: 3614

Sequence: 1 cttacatgtatgctcgtgcc.....aaaaaaaaaaaaaaaaaaaaa 2017

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 2323390

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10051902/runat_30062004_164725_20527/app_query.fasta_1.2183
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cgi -LIST=45 -DALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USAP=US10051902@cgn_1_1_13 @runat_30062004_164725_20527
-NCFU=6 -ICFU=3 -NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
------------	-------------	--------	----	-------------

1	2570	71.1	510	13	US-10-051-902-22	Sequence 22, Appl
2	2570	71.1	510	13	US-10-051-909-22	Sequence 22, Appl
3	2015	55.8	539	13	US-10-051-902-26	Sequence 26, Appl
4	2015	55.8	539	13	US-10-051-909-26	Sequence 26, Appl
5	1877.5	52.0	529	13	US-10-051-902-28	Sequence 28, Appl
6	1877.5	52.0	529	13	US-10-051-909-28	Sequence 28, Appl
7	1872.5	51.8	513	13	US-10-051-902-20	Sequence 20, Appl
8	1872.5	51.8	513	13	US-10-051-909-20	Sequence 20, Appl
9	1823	50.2	364	12	US-10-425-114-48931	Sequence 48931, A
10	1660	45.9	523	13	US-10-051-902-24	Sequence 24, Appl
11	1660	45.9	523	13	US-10-051-909-24	Sequence 24, Appl
12	1589	44.0	548	13	US-09-774-381-40	Sequence 40, Appl
13	1589	44.0	549	13	US-10-051-902-30	Sequence 30, Appl
14	1589	44.0	549	13	US-10-051-909-30	Sequence 30, Appl
15	1582.5	43.8	517	12	US-10-425-114-56035	Sequence 56035, A
16	1546.5	42.8	541	12	US-10-425-114-56733	Sequence 56733, A
17	1502.5	41.6	580	12	US-10-425-114-67036	Sequence 67036, A
18	1437.5	39.8	407	12	US-10-425-114-49353	Sequence 49353, A
19	1427	39.7	546	12	US-10-425-114-63789	Sequence 63789, A
20	1397.5	38.7	424	12	US-10-424-599-199875	Sequence 199875, A
21	1366.5	37.8	380	12	US-10-425-114-39509	Sequence 39509, A
22	1357	37.5	488	12	US-10-424-599-169603	Sequence 169603, A
23	1357	37.5	502	12	US-10-425-114-51712	Sequence 51712, A
24	1245	34.4	356	12	US-10-425-114-51926	Sequence 51926, A
25	1184	32.2	533	15	US-10-310-154-724	Sequence 724, App
26	1149.5	31.8	327	12	US-10-425-114-68399	Sequence 68399, A
27	1135	31.4	531	12	US-10-424-599-182839	Sequence 182839, A
28	1123.5	31.1	408	12	US-10-424-599-199163	Sequence 199163, A
29	994.5	27.5	281	12	US-10-425-114-50090	Sequence 50090, A
30	982.5	27.2	417	12	US-10-425-114-49121	Sequence 49121, A
31	803	22.1	475	12	US-10-425-114-41750	Sequence 41750, A
32	782.5	21.7	336	12	US-10-425-114-63429	Sequence 63429, A
33	705.5	19.5	249	12	US-10-424-599-190445	Sequence 190445, A
34	687.5	19.0	231	12	US-10-425-114-50887	Sequence 50887, A
35	682	18.9	167	13	US-10-051-902-18	Sequence 18, Appl
36	682	18.9	167	13	US-10-051-909-18	Sequence 18, Appl
37	679.5	18.8	487	9	US-09-795-693-27	Sequence 27, Appl
38	679.5	18.8	487	13	US-10-095-139-14	Sequence 14, Appl
39	679.5	18.8	487	14	US-10-156-239-27	Sequence 27, Appl
40	679.5	18.8	487	15	US-10-199-485-27	Sequence 45, Appl
41	679.5	18.8	487	15	US-10-391-399-45	Sequence 45, Appl
42	672	18.6	488	12	US-09-794-822-11	Sequence 11, Appl
43	672	18.6	488	13	US-10-094-059-4	Sequence 4, Appl
44	672	18.6	488	14	US-10-170-528-5	Sequence 5, Appl
45	672	18.6	488	14	US-10-162-012-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1

US-10-051-902-22
; Sequence 22, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT FILING DATE: 2002-01-17
; PRIORITY APPLICATION NUMBER: US/10/051,902
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)

US-10-051-902-22

Alignment Scores:

Pred. No.: 8,21e-145 Length: 510
 Score: 2570.00 Matches: 510
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 71.11% Indels: 0
 Dbs: 13 Gaps: 0

US-10-051-902A-21 (1-2017) x US-10-051-902-22 (1-510)

QY 96 ATGGTTCGGCGCGCTCCGAGAGCGCTCCGCGCGAAGAGAGGCGACGTCGCGTTC 155
 Db 1 MetAlaSerAlaAlaLeuProGluAlaValAlaProLysLysGlyAsnValArgPhe 20
 QY 156 GCCTTCGCTCGGCATCTCCCTCCATGACCTCCCTCCCTCCGCTACGATATCGG 215
 Db 21 AlaPheAlaCysAlaIleAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 40
 QY 216 GTGATGAGCGGGCGCTCGCTATCATCAAGAAGACTTCAACATCAGTACGGGAAGTG 275
 Db 41 ValMetSerGlyAlaSerLeuTyrIleLysLysAspPheAsnIleSerAspGlyLysVal 60
 QY 276 GAGTTCATGGGATACCTCTACTCGCTCATCGGCTCTTCGCGCGCGCGGCGG 335
 Db 61 GluValLeuMetGlyIleLeuAsnLeuTyrSerIleLeuGlySerPheAlaAlaGlyArg 80
 QY 336 AGTCGAGCTGATCGCGCGCGGTACACCATCGTGTTCGCGCGCTCATATTCGCGG 395
 Db 81 ThrSerAspTrpIleGlyArgArgTyrThrIleValPheAlaAlaValIlePhePheAla 100
 QY 396 GGGGSETTCATCGGTTTCGCGTCACTACGCGATGCTATGTCGCGCGCTTCGTG 455
 Db 101 Gly**PheLeuMetGlyPheAlaValAsnTyrAlaMetLeuMetPheGlyArgPheVal 120
 QY 456 GCGGCATCGCGGTGCGGTACGCGTCACTGCGCGGTGTACACCGCGCGGTGTGCG 515
 Db 121 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSer 140
 QY 516 CCGCGTTCGCGCGGTGCTCTGACGTCTGTCGCGGTGCTGAACTCGGTCGCGG 635
 Db 141 ProAlaSerAlaArgGlyPheLeuThrSerPheProGluValPheIleAsnPheGlyIle 160
 QY 576 CTGCTCGGTTAGCTCTGAACTATGCTTCTCCGCTTCCGCTGAACTCGGTCGCGG 695
 Db 161 LeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuProLeuAsnLeuGlyTyrPArg 180
 QY 636 ATCATGCTCGGCATCGCGCGCGCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
 Db 181 IleMetLeuGlyIleGlyAlaAlaProSerValLeuLeuAlaLeuMetValLeuGlyMet 200
 QY 696 CCGAGTCCCGCGTGTGCTGATGAGGAGCGCTCCGCGGAGCGCCCAAGTGTGCTG 755
 Db 201 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 220
 QY 756 GAGAGACCTCCGACACGCGGAGGAGCGCGCGGAGCGCTTGGCGGATCAAGCGCGCC 815
 Db 221 GluLysThrSerAspThrAlaGluGluAlaGluAlaGluAlaAspIleLysAlaAla 240
 QY 816 GCCGCGATCCCTGAGAGCTGACGCGGACGTGTGACCGTCCCGACGAGGAGCGGA 875
 Db 241 AlaGlyIleProGluLeuAspGlyAspValValThrValProLysArgGlySerGly 260
 QY 876 AACGAGAGCGGTTGGAAGAGCTCATCTGTCCTCCCGACCCCGCGCATCGCGGATC 935
 Db 261 AsnGluLysArgValTrpLysGluLeuLeuSerProThrProAlaMetArgArgIle 280
 QY 936 CTGCTGTCGGGATCGGCATCCACTTCTTCAGCATGCGTGGGCAATTCATCTCGTGTG 995
 Db 281 LeuLeuSerGlyIleGlyIleHisPheGlnHisAlaLeuGlyIleHisSerValVal 300
 QY 996 TTCACAGCCCTCTCGTGTTCAGAGCCCGCGGATTACGAGACGACAAACACTTCTTGGC 1055

Db 301 PheTyrSerProLeuValPheLysSerProGlyLeuThrAsnAspLysHisPheLeuGly 320
 QY 1056 ACCACTTGGCCCTTCGGTGTCCACCAAGAGGCTTTTCATCTTGTGGCGACTTTCTTCATC 1115
 Db 321 ThrThrTrpProPheGlyValThrLysArgLeuPheIleLeuAlaThrPhePheIle 340
 QY 1116 GACGCGCTCGCGCGCGCGCTGTGTGGCAGCAGCGCGGATATCTCTCCCTC 1175
 Db 341 AspGlyValGlyArgArgProLeuLeuLeuGlySerThrGlyGlyIleLeuSerLeu 360
 QY 1176 ATCGGCTCGCGCGCGGCTCACCGCTCGTGGCCACGACCCCGCGCGCAAGATACCTTGG 1235
 Db 361 IleGlyLeuGlyAlaGlyLeuThrValValGlyGlnHisProAspAlaLysIleProTrp 380
 QY 1236 GCATCGGCTTAAGCATCGCTCCACCTCGCTACGTCGCTTCTTCTCCATCGGCTT 1295
 Db 381 AlaIleGlyLeuSerIleAlaSerThrLeuAlaTyrValAlaPhePheSerIleGlyLeu 400
 QY 1296 GCGCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCGCTCCAGTCCGCGCTGGCG 1355
 Db 401 GlyProIleThrTrpValTyrSerSerGluIlePheProLeuGlnValArgAlaLeuGly 420
 QY 1356 TGCTGCTCGGCTCGCGCGCAACCGCTCACCGCGGCTCATCTCCATGACCTTCTG 1415
 Db 421 CysSerLeuGlyValAlaAlaAsnArgValThrSerGlyValIleSerMetThrPheLeu 440
 QY 1416 TGCTGCTCAAGCCATCACCATCGCGCGAGCTTCTTCTTCTTCTTCTTCTTCTTCTG 1475
 Db 441 SerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAlaAla 460
 QY 1476 CTGCTCGGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1535
 Db 461 LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu 480
 QY 1536 ATGAGCAACTGTTCGGCAGACGCGCGCGCTCGGATCAGACGAGCCAGGAG 1595
 Db 481 MetSerLysLeuPheGlyAspThrAlaAlaSerGluSerAspGluProAlaLysGlu 500
 QY 1596 AAGAAGAGGTGGAATGGCGCGCACTAAC 1625
 Db 501 LysLysLysValGluMetAlaAlaThrAsn 510

RESULT 2

US-10-051-909-22
 ; Sequence 22, Application US/10051909
 ; Publication No. US20020199217A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Helentjaris, Tim
 ; APPLICANT: Ritz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB1163 US CIP
 ; CURRENT APPLICATION NUMBER: US/10/051,909
 ; PRIOR FILING DATE: 2002-01-17
 ; PRIOR FILING DATE: 60/083,044
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 22
 ; LENGTH: 510
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (102)
 US-10-051-909-22

Alignment Scores: 8,21e-145 Length: 510
 Pred. No.: 2570.00 Matches: 510
 Score:

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	71.11%	Indels:	0
DB:	13	Gaps:	0
US-10-051-902A-21 (1-2017) x US-10-051-909-22 (1-510)			
QY	96	ATGCTTCGGCGGCTCCGGAGCGCTCGCGCGAAGAAGAGGGCAACGTCGGTTC	155
Db	1	MetAlaSerAlaAlaLeuProGluAlaValaProLysLysLysLysValaArgPhe	20
QY	156	GCTTCGGCTCGGCATCCTCGCTCCATGACCTCCATCTCCTCGGTACGATACGG	215
Db	21	AlaPheAlaCysAlaIleuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly	40
QY	216	GTGATGAGCGGGCGCTCGCTGTACATCAAGAGGACTTCAACATCAGTCAGCGGAAGTG	275
Db	41	ValMetSerGlyAlaSerLeuTyrIleLysLysAspPheAsnIleSerAspGlyLysVal	60
QY	276	GAGGTCTCATGGGCATACCTCACTCTACCTCGCTCATCGCTCTCTCGCGCGGGCGG	335
Db	61	GluValLeuMetGlyIleuAsnLeuTyrSerIleuIleGlySerPheAlaAlaGlyArg	80
QY	336	ACGTCGAGCTGATCGCGCGCGGTACACCATCGTGTTCGCGCGCTCATATCTTCGCG	395
Db	81	ThrSerAspTyrIleGlyArgArgTyrThrIleValPheAlaAlaValIlePhePheAla	100
QY	396	GGGSGTTCCTCATGGGTTCCGCGTCACACTACGCCATGCTCATGTTCGGCGGCTCGT	455
Db	101	Gly***PheLeuMetGlyPheAlaValAsnTyrAlaMetLeuMetPheGlyArgPheVal	120
QY	456	GCCGGCATCGCGTGGGTACGCGCTCATGATCGCGCGGTGTACACCGCGAGGTGCG	515
Db	121	AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSer	140
QY	516	CCGCGTCCGGCGCTGCTTCTGACGTGCTTCGGGAGGTGTTCATCACTTCGGCATC	575
Db	141	ProAlaSerAlaArgGlyPheLeuThrSerPheProGluValPheIleAsnPheGlyIle	160
QY	576	CTGCTCGGTACGTCTCAACTATGCTTTCCTCCGCTTCGCGCTGAACCTCGGTGGCG	635
Db	161	LeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuProLeuAsnLeuGlyTyrArg	180
QY	636	ATCATGCTCGGCATCGCGCGCGCTGCTGCTGCTCGCGCTCATGGTCTCGGCATG	695
Db	181	IleMetLeuGlyIleGlyAlaAlaProSerValLeuLeuAlaLeuMetValLeuGlyMet	200
QY	696	CCGAGATCGCGCGGTGGTGGTTCATGAAGGAGCGCTCGCGGACGCCAAGTGGTGGT	755
Db	201	ProGluSerProArgTyrPleuValMetCysGlyArgLeuAlaAspAlaLysValValLeu	220
QY	756	GAGAAGACCTCCGACACGGCGGAGGAGCGCCGCGAGCGCTTCGCGCATCAAGCGGCC	815
Db	221	GluLysThrSerAspThrAlaGluAlaGluAlaAlaGluArgLeuAlaAspIleLysAlaAla	240
QY	816	GCCGGCATCCTGAGGAGCTCGAGCGGACGTGTGACCTCCCGAAGAGGAGGAGCGGA	875
Db	241	AlaGlyIleProGluGluLeuAspGlyAspValValThrValProLysArgGlySerGly	260
QY	876	AACGAGAAGCGGTGTGAAGGAGCTCATCTGCTCCCGACCCCGGCATCGCGGCATC	935
Db	261	AsnGluLysArgValTyrPysGluLeuLeuSerProThrProAlaMetArgArgIle	280
QY	936	CTGCTGTCCGGATCGGCATCCACTTCTTCAGCATCGGTGGGATTCATCCTCGTCTG	995
Db	281	LeuLeuSerGlyIleGlyIleHisPhePheGlnHisAlaLeuGlyIleHisSerValVal	300
QY	996	TTCTACAGCCCTCTCGTGTTCAGAGCCCGGATTAACGACGACAAACACTTCTTGGC	1055
Db	301	PheTyrSerProLeuValPheLysSerProGlyLeuThrAsnAspLysHisPheLeuGly	320
QY	1056	ACCATTGGCGGTTCGGTGTCAACAGAGGCTTTTCACTTGTGTGGCACTTTCTTCATC	1115

Db	321	ThrThrTrpProPheGlyValThrLysArgLeuPheIleLeuLeuAlaThrPhePheIle	340
Qy	1116	GACGGCTCGGGGGCGCGCTGTCTGGGAGACACGGCGGGATAATCTCTCTCCCTC	1175
Db	341	AspGlyValGlyArgArgProLeuLeuLeuGlySerThrGlyGlyIleIleLeuSerLeu	360
Qy	1176	ATCGGCTCGGGCGCGGGCTACCGTCGTGGCCAGCACCCCGACGCCAAGATACCTTGG	1235
Db	361	IleGlyLeuGlyAlaGlyLeuThrValValGlyGlnHisProAspAlaLysIleProIrp	380
Qy	1236	GCCATCGGCTAGCATCGCTCCACCTCGCTAGCTGCGCTTCTTCTCCATCGGCTT	1295
Db	381	AlaIleGlyLeuSerIleAlaSerThrLeuAlaLysValAlaPhePheSerIleGlyLeu	400
Qy	1396	GGCCCCATACGTGGGTHACAGTCGGAGATTTCCCGTCCAGTGGCGCGCTGGGC	1355
Db	401	GlyProIleThrTrpValTyrSerSerGluIlePheProLeuGlnValArgAlaLeuGly	420
Qy	1356	TGTCGCTCGGCTCGCGCCCAACCGGTACACAGCGGGTCACTCCATGACCTTCCCTG	1415
Db	421	CysSerLeuGlyValAlaAlaAsnArgValThrSerGlyValIleSerMetThrPheLeu	440
Qy	1416	TCGCTGTCCAAAGCCATCACCATCGGGCGAGTTCTTCTCTACTCCGSCATCGCGCG	1475
Db	441	SerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAlaAla	460
Qy	1476	CTCGCTCGGCTGTCTTACACCTACCTCCCGAGACCGCGCGCGAGCTGGAGGAG	1535
Db	461	LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu	480
Qy	1536	ATGAGCAAGCTGTTCGGCGACACGGCGCGCTCGGAATCAGACGAGCAGCAAGGAG	1595
Db	481	MetSerLysLeuPheGlyAspThrAlaAlaSerGluSerAspGluProAlaLysGlu	500
Qy	1596	AAGAGAGCTGGAATGGCGGCACCTAAC	1625
Db	501	LysLysLysValGluMetAlaAlaThrAsn	510
RESULT 3			
US-10-051-902-26			
; Sequence 26, Application US/10051902			
; Publication No. US20020178468A1			
; GENERAL INFORMATION:			
; APPLICANT: Allen, Steve			
; APPLICANT: Hitz, Bill			
; APPLICANT: Kinney, Tony			
; APPLICANT: Tingey, Scott			
; TITLE OF INVENTION: Plant Sugar Transport Proteins			
; FILE REFERENCE: BB-1163			
; CURRENT APPLICATION NUMBER: US/10/051,902			
; CURRENT FILING DATE: 2002-01-17			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922			
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14			
; NUMBER OF SEQ ID NOS: 30			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 26			
; LENGTH: 539			
; TYPE: PRT			
; ORGANISM: Triticum aestivum			
US-10-051-902-26			
Alignment Scores:			
Pred. No.:		8.74e-112	Length: 539
Score:		2015.00	Matches: 408
Percent Similarity:		84.49%	Conservative: 44
Best Local Similarity:		76.26%	Mismatches: 73
Query Match:		55.76%	Indels: 10
DB:		13	Gaps: 4
US-10-051-902A-21 (1-2017) x US-10-051-902-26 (1-539)			
Qy	48	ACTGTACAGGCCAGAGCGA-----GCTCTCTCTCTCTCTGACCAACGGAG	95

Db 6 ThrGlnGlyProArgHisAsnProGlnAlaSerArgGlyAlaSerSerThr 25
 QY 96 ATGGCTTCGGCGGCTCCCGAGGCGCTCCGCGCGGAGAGAGAGGCGCAAGCTCGGTTTC 155
 Db 26 MetAspArgAlaAlaLeuProAlaAlaValGluProLysLysLysGlyAsnValArgPhe 45
 QY 156 GCCTTCGGCTGGCGCATCTCGCTCCATGACCTCCATCTCTCCGCTACGATATCGG 215
 Db 46 AlaPheAlaCysAlaIleAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 65
 QY 216 GTGATGAGCGGGCGCTGTACATCAAGAGGACTTCAACATCAGTCAGCGGAAGGTG 275
 Db 66 ValMetSerGlyAlaSerLeuTyrIleGlnLysAspLeuLysIleAsnAspThrGlnLeu 85
 QY 276 GAGGTTCATGGGCATCTGAACCTCTACTCGCTCATCGCTCTCTCGCGCGGCGGG 335
 Db 86 GluValLeuMetGlyIleLeuAsnValTyrSerLeuIleGlySerPheAlaAlaGlyArg 105
 QY 336 AGCTCGGACTCGATCGCGCGGGGTACACCATCGTGTTCGCGCGCATATCTTCGCG 395
 Db 106 ThrSerAspTrpIleGlyArgPheThrIleValPheAlaAlaValIlePhePheAla 125
 QY 396 GGGSGTTCCTCATCGGGTTCGCGCTCAACTACGCCATCGCTCATGTTCGGCGCTTCGTG 455
 Db 126 GlyAlaLeuIleMetGlyPheSerValAsnTyrAlaMetLeuMetPheGlyArgPheVal 145
 QY 456 GCGGCTACGCGGTGGGTAGCGCTCATGATCGCGCGGTGTACACCGCGAGGTGTGCG 515
 Db 146 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValAsnThrGlyValSer 165
 QY 516 CCGGCTGCGCGGTGCTCTCTGACGTCTCGTCCGCGAGGTTCATCAACTTCGCGCATC 575
 Db 166 ProAlaSerAlaArgGlyValLeuThrSerPheProGluValPheIleAsnPheGlyIle 185
 QY 576 CTGCTCGGGTACGTCTCGAATATCTTCTCCGCTTCGCGCTGAACCTCGGTGGCGG 635
 Db 186 LeuLeuGlyTyrValSerAsnPheAlaPheAlaArgLeuSerLeuArgLeuGlyTyrArg 205
 QY 636 ATCATGCTCGGATCGCGCGCGCGCTCGGTGTCTCGCGCTCATGTCTCGGATG 695
 Db 206 IleMetLeuGlyIleGlyValAlaValProSerValLeuLeuAlaPheMetValLeuGlyMet 225
 QY 696 CCGAGTCCGCGCTGTGTCTATGAAGGAGCGCTCGCGGACCGCAAGGTGTGTG 755
 Db 226 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 245
 QY 756 GAGAGACTCCGACACCGCGGAGGAGCGCGGAGCGCTCGCGGACATCAAGCGCGCC 815
 Db 246 AlaLysThrSerAspThrProGluGluAlaGluArgIleAlaAspIleLysThrAla 265
 QY 816 GCGCGCATCCCTGAGGAGCTGACCGCGACGTGTGTGACGCTCCCGCAAG--AGAGGGAGC 872
 Db 266 AlaGlyIleProLeuGlyLeuAspGlyAspValProValProLysAsnLysGlySer 285
 QY 873 GGAACGAGAGCGGGTGTGAGAGGATCATCTCTCCGAGCTCTCTCCGAGCCCGCGCATCGCGC 932
 Db 286 SerGluGluLysArgValLeuLysAspLeuIleLeuSerProThrIleAlaMetArgHis 305
 QY 933 ATCTCTGTCTCGGATCGGCATCTCACTCTCTCCAGCATGGTGTGGGATTCATCTCCGCT 992
 Db 306 IleLeuIleAlaGlyIleGlyIleHisPhePheGlnGlnSerSerGlyIleAspAlaVal 325
 QY 993 GTCTTCTACAGCCCTCTGTGTTCAGAGCCCGGATTAACAGACGAAACACATCTCTTG 1052
 Db 326 ValLeuTyrSerProLeuValPheLysSerAlaGlyIleThrGlyAspSerArgLeuArg 345
 QY 1053 GGCACCATCTGCGCTCGGTGTACCAAGAGGCTTTTCATCTGTGTGGCGACATCTCTTC 1112
 Db 346 GlyThrThrValAlaValGlyAlaThrAsnThrValPheIleLeuValAlaThrPheLeu 365
 QY 1113 ATCGACGCGCTCGGCGCGCGCTGTGTGTGGCGACGCGGCGGATATCTCTCC 1172
 Db 366 LeuAspArgIleArgArgProLeuValLeuThrSerThrGlyGlyMetLeuValSer 385

QY 1173 CTCATCGGCTCGGCGCGGCTCACCGTCTCGCGCCACCGCGCCAGCATACCT 1232
 Db 386 LeuValGlyLeuAlaThrGlyLeuThrValIleSerArgHisProAspGlnLysIleThr 405
 QY 1233 TGGGCGCATCGGCTAAGCATCGCTCCACCGCTCGCTACGTGCGCTTCTTCTCCATCGGC 1292
 Db 406 TrpAlaIleValLeuCysIlePheCysIleMetAlaTyrValAlaPhePheSerIleGly 425
 QY 1293 CTTGGCCCATCACGTGGGTGTACAGTCTGGAGATCTTCCCGCTCCAGTGGCGCGGTG 1352
 Db 426 LeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuHisValArgAlaLeu 445
 QY 1353 GCGTGTCTCGCTCGGCTCGCGCCCAACCGCTCACAGCGGCGTCTCTCCATGACCTTC 1412
 Db 446 GlyCysSerLeuGlyValAlaValAlaAsnArgLeuThrSerGlyValIleSerMetThrPhe 465
 QY 1413 CTGTCTCTCTCCAAAGCCATCACCATCGCGCGAGCTTCTCTCTATCTCGGCTATCGGC 1472
 Db 466 IleSerLeuSerLysAlaMetThrIleGlyAlaPhePheLeuPheAlaGlyIleAla 485
 QY 1473 GCGCTCGCTGGGTGTCTTCTACACCTACCTCCCGGAGACCGCGCGGCGGCGCTGGAG 1532
 Db 486 SerPheAlaTrpValPhePhePheAlaTyrLeuProGluThrArgGlyArgThrLeuGlu 505
 QY 1533 GAGATGAGCAAGCTGTTCGCGCGACACGCGC-----GCCGCTCGGAATCAGAC 1580
 Db 506 AspMetSerSerLeuPheGlyAsnThrAlaThrHisLysGlnGlyAlaAlaGluAlaAsp 525
 QY 1581 GAGCCACGACGAGAGAGAGAGAGGTGGAATGCGCGCCACTAAC 1625
 Db 526 AspAspAlaGlyGlu---LysLysValGluMetAlaAlaThrAsn 539

RESULT 4

US-10-051-909-26
 ; Sequence 26, Application US/10051909
 ; Publication No. US20020199217A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Helentjaris, Tim
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB1163 US CIP
 ; CURRENT APPLICATION NUMBER: US/10/051,909
 ; CURRENT FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: 60/083,044
 ; PRIOR FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 26
 ; LENGTH: 539
 ; TYPE: PRT
 ; ORGANISM: Triticum aestivum
 US-10-051-909-26
 Alignment Scores:
 Pred. No.: 8,74e-112 Length: 539
 Score: 2015.00 Matches: 408
 Percent Similarity: 84.49% Conservative: 44
 Best Local Similarity: 76.26% Mismatches: 73
 Query Match: 55.76% Indels: 10
 DB: 13 Gaps: 4
 US-10-051-902a-21 (1-1017) x US-10-051-909-26 (1-539)

QY 48 ACTGTACAGCGGCCAGAGCGA-----GCCCTCTCTCTCTGACACCGGAG 95
 Db 6 ThrGlnGlyProArgHisAsnProGlnAlaSerArgGlyAlaSerThr 25
 QY 96 ATGGCTTCGCGCGCTCGCGGAGCGCTCGCGCGGAGCGCGCGGCGGCGGTTTC 155


```
QY 288 GGCATAGCACTTACTCGCTCATCGGCTCCTTCGCGCGGGGGGAGCTCGGACTG 347
Db 85 GlyIleLeuSerValTyrAlaLeuIleGlySerPheLeuGlyAlaArgThrSerAspTyr 104
QY 348 ATCGCGCGGGGACACCATCGTGTTCGCGCGGTCATATTCCTCGCGGGGGTTCCTC 407
Db 105 ValGlyArgValThrValValPheAlaAlaIlePheAsnGlySerLeuLeu 124
QY 408 ATGGGGTTCGCGTCACTAGCATCTCTATGTCGCGCGCTTCGCGCGGATCGGC 467
Db 125 MetGlyPheAlaValAsnTyrAlaMetLeuMetValGlyArgPheValThrGlyIleGly 144
QY 468 GTGGGCTACCGGCTCATGATCGCGGCTGTACACCGCGGAGTGTGCGCGGCTCGCG 527
Db 145 ValGlyTyrAlaIleMetValAlaProValTyrThrProGluValSerProAlaSerAla 164
QY 528 GTGGGCTTCGAGCTGTTCCCGGAGGTGTATCAACTTCGGCATCTCTCGGGTAC 587
Db 165 ArgGlyPheLeuThrSerPheThrGluValPheIleAsnValGlyIleLeuLeuGlyTyr 184
QY 588 GTCTCGAATATGTTTCTCCGCTTCGCTGCACTCGGTCGCGCATCATGCTCGGC 647
Db 185 ValSerAsnTyrAlaPheAlaArgLeuProLeuHisLeuSerTrpArgValMetLeuGly 204
QY 648 ATCGCGCGCGCGCTCGCTGCTCGCTCATGCTCGCATGCTCGCATGCGGAGTTCGCG 707
Db 205 IleGlyAlaValProSerAlaLeuLeuAlaLeuMetValPheGlyMetProGluSerPro 224
QY 708 CGTGGCTGTATGAGGAGCGCTCGCGAGCCCAAGTGTGCTGGAGAGAGACTTC 767
Db 225 ArgTrpLeuValMetGlyArgLeuAlaAspAlaArgAlaValLeuAlaLysThrSer 244
QY 768 GACACGGCGGAGGAGCGCTGCGCGCTGCGCATCAAGCGCGCGCGCATCCCT 827
Db 245 AspThrProGluGluAlaValGluArgLeuAspGlnIleLysAlaAlaGlyIlePro 264
QY 828 GAGAGCTCGACGCGAGCTGTGACCGTCCCGAGAGAGAGAGCGGAAACGAGACGG 887
Db 265 ArgGluLeuAspGlyAspValValMetProLysThrLysGlyGlyGlnGluLysGln 284
QY 888 GTGGGAGGAGCTCATCTGCTCCCGACCGCGCATGCGCGCATCTGCTGTCGCGG 947
Db 285 ValTrpLysGluLeuIlePheSerProThrProAlaMetArgGlyLeuLeuAlaAla 304
QY 948 ATCGGCATCCACTTCTCCAGCATCGGTGGCATTCATCCGCTGCTTCTACAGCCCT 1007
Db 305 LeuGlyIleHisPhePheGlnGlnAlaThrGlySerAspSerValValLeuTyrSerPro 324
QY 1008 CTCGTGTCAAGACCGCGGATTAACGACGACAAACACTTCTTGGCACCATTCGCG 1067
Db 325 ArgValPheGlnSerAlaGlyIleThrGlyAspAsnHisLeuLeuGlyAlaThrCysAla 344
QY 1068 TTCGGTGTCAACAGAGCGTTCATCTGTGTGGCGACTTCTTCATCGAGCGGCTCGGG 1127
Db 345 MetGlyValMetLysThrLeuPheIleLeuValAlaThrPheGlnLeuAspArgValGly 364
QY 1128 CGCGCGCGCTGTGCTGGGAGACGCGCGGATATCTCTCCCTCATCGGCTCGGC 1187
Db 365 ArgArgProLeuLeuLeuThrSerThrAlaGlyMetLeuAlaCysLeuIleGlyLeuGly 384
QY 1188 GCCGGCTCACGCTCGCGCAGACCGCGCAGATACCTTGGGCCATCGGCGCTA 1247
Db 385 ThrGlyLeuThrValValGlyArgHisProAspAlaLysValProTrpAlaIleGlyLeu 404
QY 1248 AGCATCGGCTCCACCTCGCTAGCTCGCTTCTTCTCCATCGGCGCTTGGCCCATCAG 1307
Db 405 CysIleValSerIleLeuAlaTyrValSerPhePheSerIleGlyLeuGlyProLeuThr 424
QY 1308 TGGGTGTACAGTCCGAGATCTTCCGCTCAGGTCCGCGGCTGGGCTGCTCCTCGGC 1367
Db 425 SerValTyrThrSerGluValPheProLeuArgValArgAlaLeuGlyPheAlaLeuGly 444
```

```
QY 1368 GTCGCGCGCAACCGGTCACCAGCGCGTCACTCCATGACCTTCTCTGCTGCTCCAA 1427
Db 445 ThrSerCysAsnArgValThrSerAlaAlaValSerMetSerPheLeuSerLys 464
QY 1428 GCATCATCATCGCGGAGCTTCTCTACTCCGATCGCGCGCTCGCTCGGCTG 1487
Db 465 AlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyIleAlaIleIleGlyTrpIle 484
QY 1488 TTCTTCTACACTACTCTCCGAGACCGCGCGGAGCTGCGGAGATGAGCAAGCTG 1547
Db 485 PhePhePheThrPheIleProGluThrArgGlyLeuProLeuGluIleGlyLysLeu 504
QY 1548 TTCGCGGACACGCGCGCTCGGATCATCAGACGAGCAGCC--AAGAGAGAGAGAG 1604
Db 505 PheGlyMetThrAspThrAlaValGluAlaGlnAspThrAlaThrLysAspLysAlaLys 524
QY 1605 GTG 1607
Db 525 Val 525

RESULT 7
US-10-051-902-20
; Sequence 20, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-10-051-902-20

Alignment Scores:
Pred. No.: 2,62e-103 Length: 513
Score: 1872.50 Matches: 372
Percent Similarity: 84.38% Conservative: 60
Best Local Similarity: 72.66% Mismatches: 77
Query Match: 51.81% Indels: 3
DB: 13 Gaps: 2

US-10-051-902A-21 (1-2017) x US-10-051-902-20 (1-513)
QY 96 ATGGCTTCGCGCGCTGCGGAGCGCTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 155
Db 1 MetAlaSerAspGluLeuAlaLysAlaValGluProArgLysGlyAsnValLysTyr 20
QY 156 GCCTTCGCTCGGCATCTCGCTCCATGACCTCATCTCTCGCTTACGATATCGGG 215
Db 21 AlaSerIleCysAlaIleLeuAlaSerMetAlaSerValIleLeuGlyTyrAspIleGly 40
QY 216 GTGATCAGCGGGGCTGCTGTACATCAAGAAGGACTTCAACATCATGTCACGGAGGTG 275
Db 41 ValMetSerGlyAlaAlaMetTyrIleLysAspLeuAsnIleThrAspValGlnLeu 60
QY 276 GAGGTTCATCGGCGTACTGACCTTACTCTGCTCATCGGCTCTCTCGCGGGCGG 335
Db 61 GluIleIleGlyIleLeuSerLeuTyrSerLeuPheGlySerPheAlaGlyAlaArg 80
QY 336 ACCTTCGAGTTCGCGCGGCTGACATCATCTGCTTCGCGCGCTCATATCTTCGCG 395
Db 81 ThrSerAspArgIleGlyArgGluThrValValPheAlaAlaValIlePhePheVal 100
```

QY 396 GGGGGTTCCTCATGGGTTCCCGCTCAACTACGCCATGCTCATGTTCCGGCGCTTCGTG 455
Db 101 GlySerLeuLeuMetGlyPheAlaValAsnTyrGlyMetLeuMetAlaGlyArgPheVal 120
QY 456 GCGGCGATCGGCGTGGGCTACCGCTCATGATCGCGCGGCTGTACACCGCGAGGTGCG 515
Db 121 AlaGlyValGlyValGlyGlyMetIleAlaProValTyrAlaGluIleSer 140
QY 516 CCGGCGTCCGCGGCTTCCTGACGTCTTCCGAGGTGTTTCATCAACTTCGGCATC 575
Db 141 ProAlaAlaSerArgGlyPheLeuThrThrPheProGluValPheIleAsnIleGlyIle 160
QY 576 CTGCTCGGTAAGTCTCAACTATCTTTCCTCCGCTTCCGCTCAACCTCGGTGGCGC 635
Db 161 LeuLeuGlyTyrLeuSerAsnPheAlaPheAlaArgLeuProLeuHisLeuGlyTyrParg 180
QY 636 ATCATGCTCGGATCGGCGCGCGCTCCGCTGCTGCTCGGCTCATGCTCGGCATG 695
Db 181 ValMetLeuAlaIleGlyAlaValProSerGlyLeuLeuAlaLeuLeuValPheCysMet 200
QY 696 CCGGAGTCCGCGCTGCTGCTATGAGGAGCGCTCGCGGAGCGCAAGTGTGCTG 755
Db 201 ProGluSerProArgTyrLeuValLeuLysGlyArgLeuAlaAspAlaArgAlaValLeu 220
QY 756 GAGAGACCTCCGACACCGCGGAGGAGCGCGGAGCGCTGGCGGACATCAAGCGCGC 815
Db 221 GluLysThrSerAlaThrProGluGluAlaGluArgLeuAlaAspIleLysAlaAla 240
QY 816 GCGGATCCCTGAGAGCTCGACGCGGAGCTGACCGTGTGACCGTCCCGC ---AAGAGAGGAGC 872
Db 241 AlaGlyIleProLysGlyLeuAspGlyAspValValThrValProGlyLysGluGlnGly 260
QY 873 GGAACGAGAGCGGCTGTGAGAGAGCTCATCTGTCCCGACCGCGCGCATCGCGC 932
Db 261 GlyGlyGluLeuGlnValTyrLysLeuLeuSerProThrProAlaValArg 280
QY 933 ATCTGCTGTCGGGATCGGATCCACTTCTTCAGCATGCGTGGGATTCCTTCCTC 992
Db 281 IleLeuLeuSerAlaValGlyLeuHisPheGlnGlnAlaSerGlySerAspSerVal 300
QY 993 GTCTTCTACAGCCCTCTGCTGTTCAGAGCCCGGATTAAAGAACGACAAACACTTCTG 1052
Db 301 ValGlnTyrSerAlaArgLeuPheLysSerAlaGlyIleThrAspAsnLysLeuLeu 320
QY 1053 GGCACCACTTGGCTCGCTCACCAGAGGCTTTCATCTGTGTCGACATTCCTTC 1112
Db 321 GlyValThrCysAlaValGlyValThrLysThrPhePheIleLeuValAlaThrPheLeu 340
QY 1113 ATCGAGCGGCTCGGCGCGCGCTGTGCTGTGGCAGCAGCGGCGGATATCTCTCC 1172
Db 341 LeuAspArgAlaGlyArgProLeuLeuLeuIleSerThrGlyGlyMetIleValSer 360
QY 1173 CTCATCGGCTCGGCGCGGCTCACCGTCTCGGCGACACCGCGCGCAAGATACCT 1232
Db 361 LeuIleCysLeuGlySerGlyLeuThrValAlaGlyHisHisProAspThrLysValAla 380
QY 1233 TGGGCGATCGGCTTAAGCATCGCTCCACCTCGCTGCTGCTGCTTCTTCATCGCG 1292
Db 381 TrpAlaValAlaLeuCysIleAlaSerThrLeuSerTyrIleAlaPhePheSerIleGly 400
QY 1293 CTGCGCCCATCACGTGGGTACAGCTCGAGATCTTCCGCTCCAGGTGCGCGCGCTG 1352
Db 401 LeuGlyProIleThrGlyValTyrThrSerGluIlePheProLeuGlnValArgAlaLeu 420
QY 1353 GGTGCTGCTCGGCGTCCGCGCAACCGCTCACCGCGCGCTCATCTCCATCACCTTC 1412
Db 421 GlyPheAlaValGlyValAlaSerAsnArgValThrSerAlaValIleSerMetThrPhe 440
QY 1413 CTGTCGCTCAAGGCCATCACCTCGCGGAGCTTCTCTCTACTCGGCATCGCC 1472
Db 441 LeuSerLeuSerLysAlaIleThrIleGlySerPhePheLeuTyrSerGlyIleAla 460
QY 1473 GCGCTCGCTGGGTGTTCTTCTACACCTTACCTCCCGGAGACCCCGCGCGAGCTGGAG 1532

Db 461 AlaValAlaTrpValPhePheThrCysLeuProGluThrArgGlyArgThrLeuGlu 480
QY 1533 GAGATGAGCAAGCTGTTCGCGC-----GACACGGCGCGCTCGGAATCAGACGAGCCA 1586
Db 481 GluMetGlyLysLeuPheGlyMetProAspThrGlyMetAlaGluGluAlaGluAspAla 500
QY 1587 GCCAAGGAGGAAGAAGGTGGAAATGGCGGCCACT 1622
Db 501 AlaAlaLysGluLysValValGluLeuProSerSer 512
RESULT 8
US-10-051-909-20
; Sequence 20, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Kinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BBI163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-10-051-909-20
Alignment Scores:
Pred. No.: 2,62e-103 Length: 513
Score: 1872.50 Matches: 372
Percent Similarity: 84.38% Conservative: 60
Best Local Similarity: 72.66% Mismatches: 77
Query Match: 51.81% Indels: 3
DB: 13 Gaps: 2
US-10-051-902A-21 (1-2017) x US-10-051-909-20 (1-513)
QY 96 ATGGCTTCGCGCGCTGCGGAGCGCTGCGCGGAGGAGGCAAGGCAACGTCGCTTC 155
Db 1 MetAlaSerAspGluLeuAlaLysAlaValGluProArgLysGlyAsnValLysTyr 20
QY 156 GCCTTCGCTGCGCATCTCCCTCCATGACCTCCATCTCTCTCGGTACGATTCGGG 215
Db 21 AlaSerIleCysAlaIleLeuAlaSerMetAlaSerValIleLeuGlyTyrAspIleGly 40
QY 216 GTGATGAGCGGCGCTGCTGTACATCAGAGGACTTCAACATCAGTCAGCGGAGGTG 275
Db 41 ValMetSerGlyAlaAlaMetTyrIleLysLysAspLeuAsnIleThrAspValGlnLeu 60
QY 276 GAGGTTCCTCATGGGATACCTGAACCTTACTCTGCTCATCGGCTCTCTCGCGCGGCGG 335
Db 61 GluIleLeuIleGlyIleLeuSerLeuTyrSerLeuPheGlySerPheAlaGlyAlaArg 80
QY 336 AGTTCGAGTGTGATCGCGCGGCTGACCATCGCTGTCGCGCGCTCATATCTTCGCG 395
Db 81 ThrSerAspArgIleGlyArgLeuThrValValPheAlaAlaValIlePheVal 100
QY 396 GGGGSGTTCCTCATGGGTTCCCGCTCAACTACGCCATGCTCATCTTCGCGGCTTCGTG 455
Db 101 GlySerLeuLeuMetGlyPheAlaValAsnTyrGlyMetLeuMetAlaGlyArgPheVal 120
QY 456 GCGGATCGGCGTGGGCTACGCGCTCATGATCGCGCGGCTGTACACCGCGAGGTGTCG 515
Db 121 AlaGlyValGlyValGlyTyrGlyGlyMetIleAlaProValTyrThrAlaGluIleSer 140

QY 587 GTACCCGAGCAGATGTCGGAAGTTGATGAACACCTCCGGGAAACGACGTGACGAAGCCACG 528
 Db 181 ValProGluGlnAspAlaGluValAspGluHisLeuArgGluArgGlnGluAlaThr 200
 QY 527 GCCGACCCGCGCACCTCCGGCGGTGTACACCGCCGCGATCATGAGCGGTGACGCCAC 468
 Db 201 ArgArgArgArgHisLeuGlyGlyValHisArgArgAspHisGluArgValAlaHis 220
 QY 467 GCCGATGCCGCGCCGACGAGCCGCGAATGAGATGCGGTAGTGTGAGCGGAGACCCAT 408
 Db 221 AlaAspAlaGlyHisGluAlaGluHisGluHisGlyValValAspGlyGluProHis 240
 QY 407 GAGGAACCCCCCGCGAAGATATACCGCGCGCGAACAACGATGTGTACCCGCGCGCGAT 348
 Db 241 GluGluArgProArgGluGluTyrAspGlyGlyGluHisAspGlyValProProAlaAsp 260
 QY 347 CCAGTCCGACGTCCGCGCGCGCGGAGAGAGCGGATGAGCGAGTAGAGTTCAATGCC 288
 Db 261 ProValArgArgProProArgGluGluGlyAlaAspGluArgValGluValGlnTyrAla 280
 QY 287 CATGAGAACCCTCCACTTCCCGTCACTGATGTGAAGTCCTTCTGTGATGTACAGCGACG 228
 Db 281 HisGluAsnLeuHisLeuProValThrAspValGluValLeuLeuAspValGlnArgArg 300
 QY 227 CCCGTCATCAACCCGATATCGTAGCCGAGGAGGATGAGGTATGAGGCGAGATGCC 168
 Db 301 ProAlaHisHisProAspIleValAlaGluGluAspGlyGlyHisGlyGlyGluAspGly 320
 QY 167 GCAGCGAAGGCGAAGCGAGCTCCCTCTCTTCGCGCGAGCGGCTCCGCGAGCGG 108
 Db 321 AlaGlyGluGlyGluProAspValAlaLeuLeuLeuArgAspGlyLeuArgGlnArg 340
 QY 107 GCGGAAGCCATCTCCGTTGTGACAGAGGAGGAGGCTCGCTCTGGGCGGTGTACAGT 48
 Db 341 GlyGlySerHisLeuArgTrpCysArgGlyGlyGlyGlySerLeuTrpAlaValTyrSer 360
 QY 47 AGTGGCGGTGCA 36
 Db 361 SerGlyGlyArg 364

RESULT 10

US-10-051-902-24
 ; Sequence 24, Application US/10051902
 ; Publication No. US20020178468A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/10/051,902
 ; CURRENT FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 24
 ; LENGTH: 523
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 US-10-051-902-24

Alignment Scores:
 Pred. No.: 1,166-90 Length: 523
 Score: 1660.00 Matches: 331
 Percent Similarity: 75.76% Conservative: 69
 Best Local Similarity: 62.69% Mismatches: 110
 Query Match: 45.93% Indels: 18
 DB: 13 Gaps: 4

US-10-051-902A-21 (1-2017) x US-10-051-902-24 (1-523)

QY 99 GCTTCCGCGCGGTGCCGAGGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 158
 Db 12 AlaHisLysThrLeuGlnAspPheAspProProLysLysArgLysArgAsnLysTyrAla 31
 QY 159 TTGCGCTGCGCATCTCCGCTCCATGACCTCATCTCTCTCGGTACGATATCGGGTG 218
 Db 32 PheAlaCysAlaMetLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGlyVal 51
 QY 219 ATGAGCGGCGGTGCGGTACATCAAGAGGACTTCAACATCAGTCAGCGGAGGTGGAG 278
 Db 52 MetSerGlyAlaAlaIleIleLysArgAspLeuLysValSerAspGluGlnIleGlu 71
 QY 279 GTTCTCATCGGCATCACTGAACCTCTACTCGCTCATCGGCTCTCTCGCGCGCGCGGAG 338
 Db 72 IleLeuLeuGlyIleIleAsnLeuTyrSerLeuIleGlySerCysLeuAlaGlyArgThr 91
 QY 339 TCGGACTGGATCGCGCGGTACACCATCGTGTTCGCGCGCGCGGTATATCTTCGCGGG 398
 Db 92 SerAspTrpIleGlyProArgTyrThrIleValPheAlaGlyThrIlePheValGly 111
 QY 399 GSGTTCTCATCGGCGGTTCGCGCTCAACTAGCGCATGCTCTCATGTTTCGCGCGCTTC 458
 Db 112 AlaLeuLeuMetGlyPheSerProAsnTyrSerPheLeuMetPheGlyArgPheValAla 131
 QY 459 GGCATCGCGGTGGGTACCGCTCATGTCGCGCGGTGTACACCGCGCGGTGTCGCG 518
 Db 132 GlyIleGlyIleGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerPro 151
 QY 519 GCGTCGCGCGGTTCCTCTGACGTCTCCGCGAGGTTCATCAACTTCGCGCATCTCTG 578
 Db 152 AlaSerSerArgGlyPheLeuThrSerPheProGluValPheIleAsnGlyIleLeu 171
 QY 579 CTCGGTACGTCTGAACATATGCTTCTCCGCTTCGCGCTGAACTCGGGTGGCGCATC 638
 Db 172 IleGlyTyrIleSerAsnTyrAlaPheSerLysLeuThrLeuLysValGlyTrpArgMet 191
 QY 639 ATGCTCGCATCGCGCGCGGTTCGCGCTGCGCTCATGTTTCGCGCATCGCG 698
 Db 192 MetLeuGlyValGlyAlaIleProSerValLeuLeuThrValGlyValLeuAlaMetPro 211
 QY 699 GAGTCGCGCGGTGGGTGTGATGAAGGACGCTCGCGGAGCGCGCAAGTGTGTCTGAG 758
 Db 212 GluSerProArgTrpLeuValMetArgGlyArgLeuGlyGluAlaArgLysValLeuAsn 231
 QY 759 AAGACCTCCGACACGCGGAGGAGCGCGGAGCGCTGCGCGCATCAAGCCCGCGCG 818
 Db 232 LysThrSerAspSerLysGluAlaGlnLeuArgLeuAlaGluIleLysGlnAlaAla 251
 QY 819 GGCATCCCTGAGGAGCTCCGCGGACGCTGCGTGCCTCCCGAAGAGAGGAGCGGAAAC 878
 Db 252 GlyIleProGluSerCysAsnAspValValGlnValAsnLysGlnSerAsnGly--- 270
 QY 879 GAGAAGCGGTGGAAGAGCTCATCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 938
 Db 271 ---GluGlyValTrpLysGluLeuPheLeuTyrProThrProAlaIleArgHisIleVal 289
 QY 939 CTGTCGCGGATCGGATCCACTTCTCCAGCATGCGTGGGCGATTCACTCCGCTCTCTC 998
 Db 290 IleAlaAlaLeuGlyIleHisPheGlnGlnAlaSerGlyValAspAlaValValLeu 309
 QY 999 TACAGCCCTCTGTTTCAAGAGCCCCCGGATTAAAGACGACAAACACTTCTTGGGCGAC 1058
 Db 310 TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuLeuAla 329
 QY 1059 ACTTGGCCGCTCGGTGTCACCAAGAGGCTTTTCATCTGTTGCGCGCTTCTTCATCGAC 1118
 Db 330 ThrValAlaValGlyPheValLysThrValPheIleLeuAlaAlaThrPheThrLeuAsp 349
 QY 1119 GCGTCGCGCGCGCGCGCTGTTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1178
 Db 350 ArgValGlyArgArgProLeuLeuLeuSerSerValGlyGlyMetValLeuSerLeuLeu 369

```
QY 1179 GGCCTCGCGCGCGGCTCACCGTGTGGCGGACACCCGACGCAAGATACCTTGGGC 1238
Db 370 ThrLeuAlaIleSerLeuThrValIle---AspHisSerGluArgLysLeuMetTrpAla 388
QY 1239 ATCGGCTTAAGCATCGCTCCACCTCGCTACGCTACGCTCTCTCTCTCCATCGGCGCTGGC 1298
Db 389 ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly 408
QY 1299 CCATCATCGTGGTGTACAGTCGGAGATCTCCGCTCCAGGTGCGCGGCTGGGTGC 1358
Db 409 ProIleThrTpValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnGlyAla 428
QY 1359 TCGCTCGCGCGCGCGCAACCGGTCCACGAGCGGTCTCTCATGACCTTCTCTCTCG 1418
Db 429 AlaAlaGlyAlaAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448
QY 1419 CTGTCCAAGGCCATCACCATCGCGCGGAGCTTCTCTCTACTCCGCACTCGCGCGCTC 1478
Db 449 LeuThrArgAlaIleThrIleGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal 468
QY 1479 GCCTGGGTGTCTTCTACACCTACTCTCCGAGACCGCGCGCGGAGCTGGAGAGATG 1538
Db 469 GlyTrpIlePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488
QY 1539 AGCAAGCTGTTCGGCGCACACGCGCGCGCTCGGAATCAGACGACGCGCAAGGAGAG 1598
Db 489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla-----Ser 502
QY 1599 AAGAAGTGGAAATGGCGCGCACTAACTGATCAAACTAACCGCAAAATCAAAATCCTTA 1658
Db 503 LysAlaValGlu-----AsnGluAsnGlyGlnValAla 513
QY 1659 AGGGTTTCTTGCAAAACGTGTG 1682
Db 514 GlnValGlnLeuGlyThrAsnVal 521
```

RESULT 11

US-10-051-909-24

; Sequence 24, Application US/10051909

; Publication No. US20020199217A1

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Helentjaris, Tim

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: BB1163 US CIP

; CURRENT APPLICATION NUMBER: US/10/051,909

; CURRENT FILING DATE: 2002-01-17

; PRIOR APPLICATION NUMBER: 60/083,044

; PRIOR FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 24

; TYPE: PRT

; ORGANISM: Glycine max

US-10-051-909-24

Alignment Scores:

Pred. No.: 1,16e-90

Score: 1660.00

Percent Similarity: 75.76%

Best Local Similarity: 62.69%

Query Match: 45.93%

DB: 13

Length: 523

Matches: 331

Conservative: 69

Mismatches: 110

Indels: 18

Gaps: 4

US-10-051-902A-21 (1-2017) x US-10-051-909-24 (1-523)

QY 99 GCTTCGCGCGCTCGCGAGCGCTCGCGCGCGAGAGAGGCAAGTCCGGTTCGCC 158

Db 12 AlaHisLysThrLeuGlnAspPheAspProProLysLysArgLysArgAsnLysTyrAla 31

QY 1239 ATCGGCTAAGCATCGCTCCACCTCGCTACGTCGCTTCTTCTCCATCGGCTTGGC 1298
Db : : : : :
QY 389 ValGlySerSerIleAlaMetValLeuAlaValValAlaThrPheSerIleGlyAlaGly 408
Db : : : : :
QY 1299 CCATCACTGCTGCTACAGTCTCGAGATCTTCCCGCTCCAGGTCGCGCGCTGGTGC 1358
Db : : : : :
QY 409 ProlleThrTrpValTrpSerSerGluilePheProLeuArgLeuArgAlaGlnGlyAla 428
Db : : : : :
QY 1359 TCGCTCGGCTCGCGCCCAACCGCTACACGCGCGCTACTCTCCATCGCTTCTGTCG 1418
Db : : : : :
QY 429 AlaAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448
Db : : : : :
QY 1419 CTGTCCAAAGGCATCACCATCGGCGCAGCTTCTCTCTACTCGGATCGCGCGCTC 1478
Db : : : : :
QY 449 LeuThrArgAlaIleThrIleGlyAlaPhePheLeuThrCysGlyIleAlaThrVal 468
Db : : : : :
QY 1479 GCTCGGCTTCTTCTACACCTACTCTCCGAGACCCCGCGCGCTGGAGAGATG 1538
Db : : : : :
QY 469 GlyTTrpIlePhePheThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488
Db : : : : :
QY 1539 AGCAAGCTGTCGCGCACGCGCCCGCTCGGATCAGACGAGCCAGCCAGGAGAG 1598
Db : : : : :
QY 489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla-----Ser 502
Db : : : : :
QY 1599 AAGAAAGTGGAATGCGCGCCACTAACTGATCAAACTAACCGCAAAATCACAATCCTA 1658
Db : : : : :
QY 503 LysAlaValGlu-----AsnGluAsnGlyGlnValAla 513
Db : : : : :
QY 1659 AGGGTTTCTTCCAAAACGGTGTG 1682
Db : : : : :
QY 514 GlnValGlnLeuGlyThrAsnVal 521
Db : : : : :

RESULT 12

US-09-774-381-40
; Sequence 40, Application US/09774381
; Publication No. US2003008267A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Pan, Yang
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NOVEL EDIFR. MTR-1, LSP-1, TAP-1, AND PA-1 MOLECULES
; FILE REFERENCE: MNI-107CP2
; CURRENT APPLICATION NUMBER: US/09/774,381
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 08/941,354
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/010,674
; PRIOR FILING DATE: 1998-01-22
; PRIOR APPLICATION NUMBER: 60/061,149
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/014,347
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 60/061,159
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/474,151
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 09/004,206
; PRIOR FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/061,143
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/483,414
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 09/213,571
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,890
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 548
; TYPE: PRT

; ORGANISM: Beta vulgaris
US-09-774-381-40
Alignment Scores:
Pred. No.: 1,95e-86 Length: 548
Score: 1589.00 Matches: 326
Percent Similarity: 72.23% Conservative: 59
Best Local Similarity: 61.16% Mismatches: 126
Query Match: 43.97% Indels: 22
DB: 10 Gaps: 5
US-10-051-902A-21 (1-2017) x US-09-774-381-40 (1-548)
QY 78 TCCTCTGCACCCGAGAGATGCTCCGCGCGCTG-----CCGAGGCGCGTC 125
Db 9 SerAspProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 28
QY 126 GCCTCGAAGAAGAGGCAACCTCCGTTTCGCTTCGCTCGCCCATCTCCGCTCCATG 185
Db : : : : :
QY 29 ProProLysArgAsn-----LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 45
QY 186 ACCTCCATCTCTCTCGCTACGATATCGGGGTGATGAGCGGGCGCTGCTGATCATCAG 245
Db : : : : :
QY 46 ThrSerValLeuLeuGlyThrAspIleGlyValMetSerGlyAlaIleIleIleLeuLys 65
QY 246 AAGGACTTCAACATCATGTCACGGGAGGTGGAGGTTCTCATGGGATCTGAACTCTAC 305
Db 66 GluAspTrpHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 85
QY 306 TCCTCATCGCTCTCTCCGCGCGGCGAGCTCGGACTGATCGCGCGCGGCTACAC 365
Db 86 CysLeuPheGlySerPheAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThr 105
QY 366 ATCGTGTTCGCGCGCTCATATTTCTTCGCGGGGSGTCTCTCATGGGTTCCCGCTCAAC 425
Db 106 IleValLeuAlaGlyAlaIlePhePheValGlyAlaLeuLeuMetGlyPheAlaThrAsn 125
QY 426 TAGCCCATCTCATGTTCCGCGCTTGTGGCGCGCATCGCGCTGGGTACGCGCTCATG 485
Db 126 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 145
QY 486 ATCGCGCGCTGTACACCCGCGAGGTGTCGCGCGGTGCGCGCTGGCTTCTGACGTCG 545
Db 146 IleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSer 165
QY 546 TTCCCGGAGGTGTTCATCAACTTCGCGCATCTCTCGGTACGTCGGAATCATCTTTC 605
Db 166 PheProGluValPheIleAsnAlaGlyIleLeuLeuGlyTyrIleSerAsnLeuAlaPhe 185
QY 606 TCCCGCTTCGCGTGAACCTCGGTGCGCATCATCTCGGATCGCGCGCGCGCGCTCC 565
Db 186 SerSerLeuProThrHisLeuSerTrpArgPheMetLeuGlyIleGlyAlaIleProSer 205
QY 666 GTCTCTCTCGCTCATGTTCTCGCATGTCGCGAGTCCCGCGGTGGTGTGTGTATGAG 725
Db 206 IlePheLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGln 225
QY 726 GAGCGCTTCGCGAGCCCAAGGTGTGTGGAGAGACATCTCGACACGCGCGGAGAGGCC 785
Db 226 GlyArgLeuGlyAspAlaLysValLeuAsnArgIleSerAspSerProGluGluAla 245
QY 786 GCGGAGCGCTCGCGACATCAAGCGCGCGCATCTCTGAGAGGTCTCAGCGCGAC 845
Db 246 GlnLeuArgLeuSerGluIleLysGlnThrAlaGlyIleProAlaGluCysAspGluAsp 265
QY 846 GTGTGACCGTCCCCAAGAGAGGAGCGGAAACGAGAACCGGGTGTGGAAGAGGTCTATC 905
Db : : : : :
QY 266 IleTyrLysValGluLysThrLysIleLysSerGlyAsnAlaValTrpLysGluLeuPhe 285
QY 906 CTGTCCCGCGCGCGCATGCGCGCATCTCTGCTGTCGCGGATCGGATCCACATCTTTC 965
Db : : : : :
QY 286 PheAsnProThrProAlaValArgAlaValIleAlaGlyIleGlyIleHisPhePhe 305

QY 966 CAGCATGGTGGGCAATCACTCGCTCGTCTTACAGCCCTCGTGTTCAGAGAGCC 1025
Db GlnGlnAlaSerGlyIleAspAlaValLeuTyrSerProArgIlePheGlnSerAla 325
QY 1026 GAATTAACAGCAACAACACTCTTGGGACCACTTGGGACCACTTGGGCTGTCACCAAGG 1085
Db GlyIleThrAsnAlaArgLysGlnLeuLeuAlaThrValAlaValGlyValValIleThr 345
QY 1086 CTTTTCATCTTGTGGGCACTTCTTCATCAGCGGTGGGCGGCGGCGCTGTGCTG 1145
Db LeuPheIleLeuValAlaThrPheGlnLeuAspLysTyrGlyArgArgProLeuLeuLeu 365
QY 1146 GCGACAGCGGCGGATTAATCTCTCCTCATCGGCTTGGGCGGCGGCTCACCGTCTG 1205
Db ThrSerValGlyGlyMetIleAlaIleLeuThrLeuAlaMetSerLeuThrValIle 385
QY 1206 GCGCAGCACCGGACCAAGATACCTTGGGCACTGGGCACTGGGCACTGGGCACTG 1265
Db ---AspHisSerHisHisLysIleThrTrpAlaIleAlaLeuCysIleThrMetValCys 404
QY 1266 GCCTACGTCGGCTTCTTCCATCGGCTTGGGCGGCTTGGGCGGCTGACAGTCTGGAG 1325
Db AlaValValAlaSerPheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGlu 424
QY 1326 ATCTTCCCGCTCCAGGTGGGCGGCTGGGCTGGGCTGGGCTGGGCGGCAACCGGCTC 1385
Db ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 444
QY 1386 ACCAGCGGCTCATCTCCATGACCTTCTGTCGCTGTCGAGGCGCATCACCATCGGCGG 1445
Db ValSerGlyValIleSerIlePhePheLeuProLeuSerHisLysIleThrThrGlyGly 464
QY 1446 AGCTTCTTCTTCTTCCGCGATCGCGCTCGCTGCGTGGGTGTCTTCTTACACTACTC 1505
Db AlaPhePheLeuPheGlyGlyIleAlaIleAlaTrpPhePhePheLeuThrPheLeu 484
QY 1506 CCGGAGACCGCGCGGACCTCGGAGAGATGAGCAAGCTGTTCGGCGAC----- 1556
Db ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTrp 504
QY 1557 -----ACGGCGCGCGCTCGGAATCAGACGAGCGAGCC----- 1589
Db ArgGluSerPheProGlyAsnLysSerAsnAsnAspGluAsnSerThrArgLysGlnSer 524
QY 1590 -----AAGGAGAGAGAAGTGGAATGCGCGCACT 1622
Db AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThr 537

RESULT 13
; Sequence 30, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ IDS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-10-051-902-30
Alignment Scores:
Pred. No.: 1.95e-86 Length: 549

Score: 1589.00 Matches: 326
Percent Similarity: 72.23% Conservative: 59
Best Local Similarity: 61.16% Mismatches: 126
Query Match: 43.97% Indels: 22
DB: 13 Gaps: 5
US-10-051-902A-21 (1-2017) x US-10-051-902-30 (1-549)
QY 78 TCCTTCGACACCGGAGATGGTTCGCGCGCGCTG-----CCGGAGGCGGTC 125
Db 10 SerAspProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 29
QY 126 GCGCCGAAAGAGGCAACGTCCGTTCCGCTTCGCTCGCCATCTCTCGCTCCATG 185
Db 30 ProProLysArgAsn-----LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 46
QY 186 ACCTTCATCTCTCGCTACGATATCGGGGTGATGAGCGGCGGCTGCTGTACATCAAG 245
Db 47 ThrSerValLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaIleIleTyrLeuLys 66
QY 246 AAGGACTTCAACATCATGACGGGAAGGTGGAGTTCTCATGGGCATATGAACTCTAC 305
Db 67 GluAspTrpHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 86
QY 306 TCCTTCATCGGCTCTTCGCGCGGCGGAGTCCGACTGGATCGCGCGGTACACC 365
Db 87 CysLeuPheGlySerPheAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThr 106
QY 366 ATCGTGTTCGCGCGCTCATATTCCTTCGCGGCGGSGTTCCTCATGGGTTTCGCGCTCAAC 425
Db 107 IleValLeuAlaGlyAlaIlePhePheValGlyAlaLeuLeuMetGlyPheAlaThrAsn 126
QY 426 TAGCCCATCTCATGTTCCGCGCTTCGTCGGCGGCGCATCGCGTGGGCTACGCGCTCATG 485
Db 127 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 146
QY 486 ATCGCGCGGCTGACACCGCGGAGGTGTCGCGCGGTGCGCGCTGCGCTCTCTGAGTCG 545
Db 147 IleAlaProValTyrThrAlaGluValSerProAlaSerArgGlyPheLeuThrSer 166
QY 546 TTCCCGGAGGTGTTCATCAACTTCGCGATCTCTCTCGGTAGCTCTCGAACTATGCTTC 605
Db 167 PheProGluValPheIleAsnAlaGlyIleLeuLeuGlyTyrIleSerAsnLeuAlaPhe 186
QY 606 TCCGCTTCGCGCTGAACCTCGGTCGCGCATCTCTCGGCATCGCGCGGCGCGGCTCC 665
Db 187 SerSerLeuProThrHisLeuSerTrpArgPheMetLeuGlyIleGlyAlaIleProSer 206
QY 666 GTGCTCTCGCGCTCATGTGCTCGGCATCGCGGAGTCCGCGGCTGCGTGTGCTCATGAAG 725
Db 207 IlePheLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGln 226
QY 726 GAGCGCTCGCGACCCCAAGGTGCTGCTGGAGAGACCTCCGACACGCGGAGGAGGCGC 785
Db 227 GlyArgLeuGlyAspAlaLysLysValLeuAsnArgIleSerAspSerProGluGluAla 246
QY 786 GCGGAGCGCTGCGGACATCAAGCGCGCGGCGGCTCCCTGAGGAGCTCGACGCGGAC 845
Db 247 GlnLeuArgLeuSerGlnIleLysGlnThrAlaGlyIleProAlaGlyCysaspGluAsp 266
QY 846 GTGTCACCGTCCCCAAGAGAGGAGCGGAAACGAGAACGCGGTGTGGAAGAGCTCATC 905
Db 267 IleTyrLysValGluLysThrLysIleLysSerGlyAsnAlaValTrpLysGluLeuPhe 286
QY 906 CTGTCGCGACCGCGGCTTCGCGGCGATCTCTCTCGGATCGGCTCCGCTTCCTTTC 965
Db 287 PheAsnProThrProAlaValArgAlaValIleAlaGlyIleGlyIleHisPhePhe 306
QY 966 CAGCATGCGTGGGCAATTCACCTCGCTCTTCTACAGCCCTCTCTGTGTTCAGAGCCGC 1025
Db 307 GlnGlnAlaSerGlyIleAspAlaValValLeuTyrSerProArgIlePheGlnSerAla 326
QY 1026 GATTAAACGACGACAAACACTTCTTGGGCAACACTTTCGCGCTTCGCTGCTACCAAGAGG 1085

Db 347 LeuPheIleLeuValAlaThrPheGlnLeuAspLysTyrGlyArgArgProLeuLeuLeu 366
QY 1146 GGCAGCAGCGCGGAGTAATCTCTCTCCATCGCTCGCGCGCGCGCGCTCACCGTGC 1205
Db 367 ThrSerValGlyGlyMetIleAlaIleLeuThrLeuAlaMetSerLeuThrValIle 386
QY 1206 GGCCAGCAGCCGCGGAGCAATACCTTGGGCCATCGCGCTCAAGCATCGCTCCACCTC 1265
Db 387 ---AspHisSerHisLysIleThrTrpAlaIleAlaLeuCysIleThrMetValCys 405
QY 1266 GCCTACGCTCTCTCTCCATCGCTTGGCCCATCGCTCGCGCGCGCTCACCGTGCAG 1325
Db 406 AlaValAlaAlaSerPheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGlu 425
QY 1326 ATCTTCCGCTCCAGGTCGCGCGCTCGCTCGCTCGCGCGCGCTCGCGCGCGCTCG 1385
Db 426 ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 445
QY 1386 ACCAGCGGCTCATCTCATGACCTTCTCGCTCGCTCGCTCGCGCGCTCACCGTGC 1445
Db 446 ValSerGlyValIleSerIlePhePheLeuProLeuSerHisLysIleThrThrGlyGly 465
QY 1446 AGCTTCTCTCTACTCGGCTCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1505
Db 466 AlaPhePheLeuPheGlyGlyAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 485
QY 1506 CCGAGAGCCGCGCGCGAGCTCGAGAGATGAGCAAGTGTTCGCGGAC----- 1556
Db 486 ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTyr 505
QY 1557 -----ACGCGCGCGCGCTCGGAATCAGACGAGCGCGC----- 1589
Db 506 ArgGluSerPheProGlyAsnLysSerAsnAspGluAsnSerThrArgLysGlnSer 525
QY 1590 -----AAGGAGAGAGAGAGTGGAAATGCGCGCGCT 1622
Db 526 AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThr 538

RESULT 15
US-10-425-114-56035
; Sequence 56035, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56035
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701204606_FLI.pep
US-10-425-114-56035

Alignment Scores:
Pred. No.: 4,73e-86 Length: 517
Score: 1582.50 Matches: 305
Percent Similarity: 77.53% Conservative: 78
Best Local Similarity: 61.74% Mismatches: 108
Query Match: 43.79% Indels: 3
DB: 12 Gaps: 2

US-10-051-902A-21 (1-2017) x US-10-425-114-56035 (1-517)

QY 132 AAGAAGAGGGCAAGTTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 191
Db 16 LysLysProArgArgAsnLysTyrAlaPheAlaCysAlaIleLeuAlaSerMetThrSer 35
QY 192 ATCTCTCTCGCTCGATATCGGGGTGATGACCGGGGGGTGCTGCTGATCATCAAGAGGAC 251
Db 36 IleLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaAlaLeuTyrIleGlnArgAsp 55
QY 252 TTAACATCATGACGAGGAGGTTCATCGGGCATACTGAACCTCTACTCGCTC 311
Db 56 LeuLysValSerAspValGlnIleGluIleLeuAsnGlyIleIleAsnLeuTyrSerPro 75
QY 312 ATCGGCTCTCTCGCGCGCGGAGCGTTCGACTGATCGCGCGCGGCTGACACCTCGT 371
Db 76 ValGlySerPheIleAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrIleVal 95
QY 372 TTCGCGCGCTCATATTCTTCGCGGGGGTTCCTCATGGGTTCCTCGCTCAACTACGCG 431
Db 96 LeuAlaGlyAlaIlePheValGlyAlaIleLeuMetGlyPheSerProAsnTyrAla 115
QY 432 ATGCTCATCTTCGCGCGCTTCGTCGCGCATCGCGGTGGGTACGCGCTCATGATCGCG 491
Db 116 PheLeuMetPheGlyArgPhePheAlaGlyValGlyIleGlyPheAlaPheLeuIleAla 135
QY 492 CCGGTGTACACCGCGAGGTGTCGCGCGCTCGCGCGGTTCCTGACGTCTCGCTCGCT 551
Db 136 ProValTyrThrSerGluIleSerProSerSerArgGlyPheLeuThrSerLeuPro 155
QY 552 GAGGTGTTTCATCAACTTCGCGCATCTGCTCGGGTACGTCGAACTATGCTTCTCCCGC 611
Db 156 GluValPheLeuAsnGlyGlyIleLeuIleGlyTyrIleSerAsnTyrGlyPheSerLys 175
QY 612 TTCGCGCTGAACCTCGGTGGGCGATCATGTCGCGCATCGCGCGCGCTCGCTCGCT 671
Db 176 LeuAlaLeuArgLeuGlyTrpArgLeuMetLeuGlyValGlyAlaIleProSerIleLeu 195
QY 672 CTCGCGCTCATGCTCGCGCATCGCGAGTTCGCGGTGCTGCTGATGAAGGACGCG 731
Db 196 IleGlyValAlaValLeuAlaMetProGluSerProArgTrpLeuValAlaLysGlyArg 215
QY 732 CTCGCGGACGCAAGTGTGTGAGAGAGACTCGACACGCGGAGAGAGCGCGCGGAG 791
Db 216 LeuGlyGluAlaLysArgValLeuTyrLysIleSerGluSerGluGluAlaArgLeu 235
QY 792 CGCTGCGCGCATCAAGCGCGCGCGCATCTCGAGAGCTCGAGAGCTCGAGCGCGCTG 851
Db 236 ArgLeuAlaAspIleLysAspThrAlaGlyIleProGlnAspCysAspAspValVal 255
QY 852 ACCGTCCTCCAGAGAGGAGCGGAGAAACGAGAGCGGTGTGGAAGAGCTCATCTGTC 911
Db 256 LeuValSerLysGlnThrHisGlyHis-----GlyValTrpArgGluLeuPheLeuHis 273
QY 912 CGACCGCGCGCATCGCGCGATCTGTCGCGGATCGGATCCACTTCTTCAGCAT 971
Db 274 ProThrProAlaValArgHisIlePheIleAlaSerLeuGlyIleHisPhePheAlaGln 293
QY 972 GGTGGGCGATTCACTCGCTGCTCTACAGCCCTCTGTTCAAGAGCGCGCGGATTA 1031
Db 294 AlaThrGlyIleAspAlaValValLeuTyrSerProArgIlePheGluLysAlaGlyIle 313
QY 1032 ACGAAGCAAAACATTTTGGGACCATCTGCGCGCTCGGTGTCACCAAGAGGTTTTC 1091
Db 314 LysSerAspAsnTyrArgLeuLeuAlaThrValAlaValGlyPheValTyrValSer 333
QY 1092 ATCTTGTGGCGACTTCTTCATCGCGCGCTCGCGCGCGCGCTGCTGCTGGCGAGC 1151
Db 334 IleLeuValAlaThrPhePheLeuAspArgAlaGlyArgValLeuLeuLeuCysSer 353
QY 1152 ACGGGCGGATATCTCTCTCCCTCATCGCGCTCGCGCGCGCGCGCTCACCGTCTCGCG 1211
Db 354 ValSerGlyLeuIleLeuSerLeuLeuThrLeuGlyLeuSerLeuThrValVal---Asp 372

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:37:56 ; Search time 100 Seconds

(without alignments)
11397.966 Million cell updates/sec

Title: US-10-051-902A-21

Perfect score: 3614
Sequence: 1 ctctacatgaagctctgccc.....aaaaaaaaaaaaaaaaaaaaa 2017

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p.model -DBV=xlp
-Q=/cgn21/USFTO_spool_P/US10051902/runat_30062004_164722_20393/app_query.fasta_1.2183
-DB=A_Geneseq_29Jan04 -Q=FASTA -SUFFIX=rag -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10051902 @CGN 1.1 81 @runat_30062004_164722_20393 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPEXT=6
-FGAPEXT=7 -YGAPOPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_29Jan04:.*
1: Geneseqp1980s.*
2: Geneseqp1980s.*
3: Geneseqp2000s.*
4: Geneseqp2000s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2570	71.1	510	5 AAU97211	Aau97211 Rice Beta
2	2570	71.1	510	6 ABU08336	Abu08336 Rice suga
3	2015	55.8	539	5 AAU97213	Aau97213 Wheat sug
4	2015	55.8	539	6 ABU08338	Abu08338 Wheat sug
5	1877.5	52.0	529	5 AAU97214	Aau97214 Wheat sug
6	1877.5	52.0	529	6 ABU08339	Abu08339 Wheat sug
7	1872.5	51.8	513	5 AAU97210	Aau97210 Corn Beta
8	1872.5	51.8	513	6 ABU08335	Abu08335 Corn suga
9	1660	45.9	523	5 AAU97212	Aau97212 Soybean B
10	1660	45.9	523	6 ABU08337	Abu08337 Soybean s

11	1513	41.9	513	5 ABB09681	Abb09681 Amino aci
12	1385.5	38.3	491	3 AAG32072	Aag32072 Arabidops
13	1385.5	38.3	508	3 AAG32071	Aag32071 Arabidops
14	1303.5	36.1	466	3 AAG32073	Aag32073 Arabidops
15	827.5	22.9	333	6 ADA48320	Ada48320 Rice prot
16	682	18.9	167	5 AAU97209	Aau97209 Portion o
17	682	18.9	167	6 ABU08334	Abu08334 Corn suga
18	679.5	18.8	487	6 ABG73334	Abg73334 Consensus
19	672	18.6	488	6 ABP98504	Abp98504 PFAM cons
20	672	18.6	488	6 AAB35306	Aae35306 Human sug
21	672	18.6	488	7 ADD22918	Add22918 Human sug
22	658.5	18.2	629	4 AAB66935	Aab66935 Human GLU
23	658.5	18.2	629	4 AAB66940	Aab66940 GLUTX2 co
24	653.5	18.1	648	5 ABG61548	Abg61548 Human tra
25	649	18.0	446	5 ABU43281	Abu43281 Protein e
26	639.5	17.7	618	4 AAB66936	Aab66936 Rat GLUTX
27	629.5	17.4	582	3 AAG29528	Aag29528 Arabidops
28	628.5	17.4	555	3 AAG29529	Aag29529 Arabidops
29	594	16.4	517	5 ABP66239	Abp66239 Bifidobac
30	591	16.4	478	3 AAG15416	Aag15416 Arabidops
31	591	16.4	493	3 AAG15415	Aag15415 Arabidops
32	591	16.4	456	3 AAG15414	Aag15414 Arabidops
33	584.5	16.2	450	6 ABU33853	Abu33853 Protein e
34	582	16.1	491	4 AAG89949	Aag89949 C Glutam
35	580	16.0	491	3 AAB12594	Aab12594 Brevibact
36	579.5	16.0	472	5 ABP52164	Abp52164 E. coli a
37	579.5	16.0	472	7 ADD37502	Add37502 E. coli a
38	576	15.9	584	5 ABB05603	Abb05603 Yeast ino
39	572.5	15.8	464	5 ABP52163	Abp52163 E. coli g
40	572.5	15.8	464	6 ABU14994	Abu14994 Protein e
41	572.5	15.8	464	7 ADD37501	Add37501 E. coli g
42	571.5	15.8	465	6 ABU28442	Abu28442 Protein e
43	571	15.8	464	6 ABU48384	Abu48384 Protein e
44	571	15.8	464	6 ABU46992	Abu46992 Protein e
45	569	15.7	453	6 ABU31194	Abu31194 Protein e

ALIGNMENTS

RESULT 1

AAU97211
ID AAU97211 standard; protein; 510 AA.

XX
AC AAU97211;

DT 27-AUG-2002 (first entry)

DE Rice Beta vulgaris-like sugar transport protein.

KW Rice; Beta vulgaris-like sugar transport protein; carbohydrate transport;
grain filling; annual field crop; plant.

XX Oryza sativa.

XX Key Location/Qualifiers
FH Key Misc-difference 102
FT Misc-difference /label= Unknown

XX US6383776-B1.

PD 07-MAY-2002.

PF 14-APR-1999; 99US-00291922.

PR 24-APR-1998; 98US-0083044P.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

DR WPI; 2002-453364/48.

DR N-PSDB; ABK51972.

XX

PT New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution.
 XX
 PS Example 4; Fig 2; 54pp; English.
 XX
 CC The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
 CC or Beta vulgaris-like sugar transport protein. The polynucleotide
 CC sequences are useful for altering the level of sugar transport proteins
 CC in plants, i.e. for control of carbohydrate transport and distribution in
 CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
 CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
 CC sugar transport. The polynucleotide sequences can also be used to isolate
 CC cDNA sequences and genes that encode homologues of the new proteins. The
 CC present sequence represents a rice Beta vulgaris-like sugar transport
 CC protein
 XX
 SQ Sequence 510 AA;

Alignment Scores:
 Pred. No.: 9, 06e-186 Length: 510
 Score: 2570.00 Matches: 510
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 71.11% Indels: 0
 DB: 5 Gaps: 0

US-10-051-902a-21 (1-2017) x AAU97211 (1-510)

QY 96 ATGGCTCCGCGCGCTGCGGAGCGCTGCGCGGAGAGAGGCGCAAGCTCCGGTTC 155
 Db 1 MetAlaSerAlaAlaLeuProGluAlaValAlaProLysLysLysGlyAsnValArgPhe 20

QY 156 GCCTTCGCGTCCGCGCATCTCGCTCCATGACCTCCATCTCCCTCCGCTACGATATCGGG 215
 Db 21 AlaPheAlaCysAlaAlaLeuAlaSerMetThrSerIleuLeuGlyTyrAspIleGly 40

QY 216 GTGATGACGGGGCGCTCGCTATACATCAAGAAGCACTCAACATCACTACGGGAAGGTG 275
 Db 41 ValMetSerGlyAlaSerLeuTyrIleLysLysAspPheAsnIleSerAspGlyLysVal 60

QY 276 GAGGTTCTATGGGATCACTAACTCTACTCGCTCATCGCTCTTTCGGCGGGGCGG 335
 Db 61 GluValLeuMetGlyIleLeuAsnLeuTyrSerLeuIleGlySerPheAlaAlaGlyArg 80

QY 336 ACGTCGGAATCGGCGCGGTACACCATCGTGTTCGCGCGGTCAATATTCTTCGG 395
 Db 81 ThrSerAspTrpIleGlyArgTyrThrIleValPheAlaAlaValIlePhePheAla 100

QY 396 GGGSGTTCATCGGGGTTCCGGCTCACTACGCGATGCTATGTTTCGGCGCGCTTCGTG 455
 Db 101 Gly***PheLeuMetGlyPheAlaValAsnTyrAlaMetLeuMetPheGlyArgPheVal 120

QY 456 GCCGCACTCGCGGTGCTACCGCTCATGATCGCGCGGTGTACACCGCGGAGGTGTCG 515
 Db 121 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSer 140

QY 516 CCGCGCTCGGCGCGGTTCCTGACGCTGTTCCGCGAGGTGTTTCATCACTTCGGCATC 575
 Db 141 ProAlaSerAlaArgGlyPheLeuThrSerPheProGluValPheIleAsnPheGlyIle 160

QY 576 CTGCTCGGATCGCTCAACTATGCTTCTCCGCTTCCGCTGGAACCTCGGTTGCGCG 635
 Db 161 LeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuProLeuAsnLeuGlyTrpArg 180

QY 636 ATCATGCTCGCATCGGCGCGCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 695
 Db 181 IleMetLeuGlyIleGlyAlaAlaProSerValLeuLeuAlaLeuMetValLeuGlyMet 200

QY 696 CCGAGTCCGCGCGGTGCTGCTCATGAGGAGCGCTCGCGGACGCCAAGGTGCTGCTG 755
 Db 201 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValLeu 220

QY 756 GAGAAAGCTCCGACACGCGGAGGAGCGCGGAGCGCTCGCGCATCAAGGCGCGCC 815
 Db 221 GluLysThrSerAspThrAlaGluAlaGluAlaGluAlaGluAlaAspIleLysAlaAla 240

QY 816 GCCGCATCCCTAGGAGCTCGACGCGACGCTGTGTGACCGTCCCGCCCAAGAGGAGCGGA 875
 Db 241 AlaGlyIleProGluLeuLeuAspGlyAspValValThrValProLysArgGlySerGly 260

QY 876 AACGAGAGCGGCTGGAAGAGCTCATCTCCCGACCGCGCGCATCGCGCGCATC 935
 Db 261 AsnGluLysArgValTrpLysGluLeuLeuLeuSerProThrProAlaMetArgArgIle 280

QY 936 CTGCTGTCCGGATCGGCATCCACTTCTTCAGCATGCTGTGGCATTCATCTCGTCTC 995
 Db 281 LeuLeuSerGlyIleGlyIleHisPheGlnHisAlaLeuGlyIleHisSerValVal 300

QY 996 TTCTACAGCCCTCTGCTTCAAGAGCCCGGATTAACGACGACAAACACTTCTTGGGC 1055
 Db 301 PheTyrSerProLeuValPheLysSerProGlyLeuThrAsnAspLysHisPheLeuGly 320

QY 1056 ACCACTTGGCGCTTCGCTGTCTACCAAGAGGCTTTTTCATCTTGTGGCGACTTCTTCATC 1115
 Db 321 ThrThrTrpProPheGlyValThrLysArgLeuPheIleLeuAlaThrPhePheIle 340

QY 1116 GACGCGTCCGCGCGCGCGCTGTGTGGGACGACGCGCGGATTAATCTCTCCCTC 1175
 Db 341 AspGlyValGlyArgProLeuLeuLeuGlySerThrGlyGlyIleIleLeuSerLeu 360

QY 1176 ATCGGCTCGGCGCGGCTCACCTCGTCCGCGACACCGCGCGCGCAAGACTCTTGG 1235
 Db 361 IleGlyLeuGlyAlaGlyLeuThrValValGlyGlnHisProAspAlaLysIleProTrp 380

QY 1236 GCCATCGCGCTAAGCATCGCTCCACCTCGCTACGTCCGCTTCTTCTCCATCGGCTT 1295
 Db 381 AlaIleGlyLeuSerIleAlaSerThrLeuAlaTyrValAlaPhePheSerIleGlyLeu 400

QY 1296 GCGCCCATCATCGTGGGTGTACAGTCGCGAGATCTCCGCTCCAGGTCCGCGCGTGGC 1355
 Db 401 GlyProIleThrTrpValTyrSerSerGluIlePheProLeuGlnValArgAlaLeuGly 420

QY 1356 TCCTCGCTCGGCGTCCGCGCAACCGCTCACCGCGCTCACCGCGCTCATCTCCATCACTTCTG 1415
 Db 421 CysSerLeuGlyValAlaAlaAsnArgValThrSerGlyValIleSerMetThrPheLeu 440

QY 1416 TCGCTGTCCAGGCGCATCACATCGCGCGAGCTTCTCTCTACTCCGGATCGCGCG 1475
 Db 441 SerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAlaAla 460

QY 1476 CTCGCTCGGCTGTTCTTCTACACCTACCTCCGAGACCGCGCGCGGAGAG 1535
 Db 461 LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu 480

QY 1536 ATGAGCAAGCTTTCGCGACACGCGCGCTCGGATTCGGAATCAGACGCGCGAGAG 1595
 Db 481 MetSerLysLeuPheGlyAspThrAlaAlaAlaSerGluSerAspGluProAlaLysGlu 500

QY 1596 AAGAGAGGTCGGAATGCGCGCACTAAC 1625
 Db 501 LysLysLysValGluMetAlaAlaThrAsn 510

RESULT 2
 ABU08336
 ID ABU08336 standard; protein; 510 AA.
 XX
 AC ABU08336;
 XX
 DT 29-MAY-2003 (first entry)
 XX
 DE Rice sugar transport protein #3.
 XX
 KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant.

Db 461 LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu 480
 QY 1536 ATGAGCAAGCTGTCGGCGACACGGCGCGCCCTCGGAATCAGACGAGCCAGCAGGAG 1595
 Db 481 MetSerLysLeuPheGlyAspThrAlaAlaAlaSerGluSerAspGluProAlaLysGlu 500
 QY 1596 AAGAGAGAGGTGGAAATGGCGCCCACTAAC 1625
 Db 501 LysLysLysValGluMetAlaAlaThrAsn 510
 RESULT 3
 AAU97213
 ID AAU97213 standard; protein; 539 AA.
 AC AAU97213;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Wheat sugar transport protein encoded by wlk8.pk0001.all.
 XX
 KW Wheat; Beta vulgaris-like sugar transport protein;
 KW carbohydrate transport; grain filling; annual field crop; plant.
 XX
 OS Triticum aestivum.
 XX
 PN US6383776-B1.
 XX
 PD 07-MAY-2002.
 XX
 PF 14-APR-1999; 99US-00291922.
 XX
 PR 24-APR-1998; 98US-0083044P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 DR WPI: 2002-453364/48.
 DR N-PSDB; ABK51974.
 XX
 XX New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution.
 PT
 XX
 PS Example 4; Fig 2; 54pp; English.
 XX
 CC The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
 CC or Beta vulgaris-like sugar transport protein. The polynucleotide
 CC sequences are useful for altering the level of sugar transport proteins
 CC in plants, i.e. for control of carbohydrate transport and distribution in
 CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
 CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
 CC sugar transport. The polynucleotide sequences can also be used to isolate
 CC cDNA sequences and genes that encode homologues of the new proteins. The
 CC present sequence represents a wheat Beta vulgaris-like sugar transport
 CC protein
 XX
 SQ Sequence 539 AA;
 Alignment Scores:
 Pred. No.: 9,75e-144 Length: 539
 Score: 2015.00 Matches: 408
 Percent Similarity: 84.49% Conservative: 44
 Best Local Similarity: 76.26% Mismatches: 73
 Query Match: 55.76% Indels: 10
 Ds: 5 Gaps: 4
 US-10-051-902a-21 (1-2017) x AAU97213 (1-539)
 QY 48 ACTGTACACGGCCAGACGCA-----GCCTCTCTCTCTCTGTGACACCGGAG 95
 Db 6 ThrGlnGlyGlyProArgArgHisAsnProGlnAlaAlaSerArgGlyAlaSerThr 25

QY 96 ATGGTTTCGCGCGCTCGCGAGCGCGCTCGCGCCGAAGAAGAGGGCAACGTTCGGTTC 155
 Db 26 MetAspArgAlaAlaLeuProAlaAlaValGluProLysLysLysLysValArgPhe 45
 QY 156 GCCTTCGCTCGCGCATCTCGCTCCATGACCTCCATCCCTCGCTCGGTACGATACGG 215
 Db 46 AlaPheAlaCysAlaAlaLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 65
 QY 216 GTGATGAGCGGGCGCTCGCTGTACATCAAGAAGACTTCAACATCAGTACGGGAAGTG 275
 Db 66 ValMetSerGlyAlaSerLeuTyrIleGlnLysAspLeuLysIleAsnAspThrGlnLeu 85
 QY 276 GAGGTTCATCGGCATACCTGACCTCTACTCGCTCATCGCTCGCTCGCTCGCGCGGG 335
 Db 86 GluValLeuMetGlyIleLeuAsnValTyrSerLeuIleGlySerPheAlaAlaGlyArg 105
 QY 336 ACGTCGACTGATCGCGCGCGGTACACCATCTGTGTTCGCGCGCTCATATCTTCGCG 395
 Db 106 ThrSerAspTrpIleGlyArgPheThrIleValPheAlaAlaValIlePhePheAla 125
 QY 396 GGGGSGTTCCTCATGGGTTCCCTCAACTACGCCATGCTCATGTTCGGCGCTTCGTG 455
 Db 126 GlyAlaLeuIleMetGlyPheSerValAsnTyrAlaMetLeuMetPheGlyArgPheVal 145
 QY 456 GCGGCATCGCGGTGGCTACCGCTCATGTCGCGCGGTGTACACCGCGCGAGTGTG 515
 Db 146 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValAsnThrGlyGluValSer 165
 QY 516 CCGCGTCGCGCGCTGCTTCAGCTGTCGCGAGGTTCATCACTTCGGGTATC 575
 Db 166 ProAlaSerAlaArgGlyValLeuThrSerPheProGluValPheIleAsnPheGlyIle 185
 QY 576 CTGCTCGGGTAGCTCTCGAACTATCTTTCCTCGCTTCGCGCTCAACTCGGGTGGCG 635
 Db 186 LeuLeuGlyTyrValSerAsnPheAlaPheAlaArgLeuSerLeuArgLeuGlyTrpArg 205
 QY 636 ATCATGCTCGGCATCGCGCGCGCTCGCTCGCTCGCTCGCTCATGCTCGGTATC 695
 Db 206 IleMetLeuGlyIleGlyAlaValProSerValLeuLeuAlaPheMetValLeuGlyMet 225
 QY 696 CCGGAGTGCCTCGGTGCTGTATGAAGGACGCTCGCGGAGCCCTCGGAGCGCAAGTGTG 755
 Db 226 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 245
 QY 756 GAGAGACCTCCGACACACGCGGAGGCGCGGAGCGCTGCGCGACATCAAGCGCGCC 815
 Db 246 AlaLysThrSerAspThrProGluGluAlaAlaGluArgIleAlaAspIleLysThrAla 265
 QY 816 GCGGCGATCCCTGAGAGCTCGACGCGGACGTGTGTGACCGCTCCCAAG---AGAGGGAGC 872
 Db 266 AlaGlyIleProLeuGlyLeuAspGlyAspValProValProLysAsnLysGlySer 285
 QY 873 GGAACGAGAACGGGTGTGGAAGAGCTCATCTGTCCCGACCCCGCCATCGGCGCGC 932
 Db 286 SerGluGluLysArgValLeuLysAspLeuLeuSerProThrIleAlaMetArgHis 305
 QY 933 ATCTGCTGTCCGGATCGGCATCCACTTCTTCCAGCATGTGCGGTTCGGGATCATCTCGTC 992
 Db 306 IleLeuIleAlaGlyIleGlyIleHisPhePheGlnGlnSerSerGlyIleAspAlaVal 325
 QY 993 GTCTTCTACAGCCCTCTCGTGTTCAGAGCCCGCGGATTAAGAGCAACACACTTCITG 1052
 Db 326 ValLeuTyrSerProLeuValPheLysSerAlaGlyIleThrGlyAspSerArgLeuArg 345
 QY 1053 GGCACCACTTGGCGCTCGGTGTCCACCAAGAGCGCTTTTCATCTTGTTCGCGACTTCTTC 1112
 Db 346 GlyThrThrValAlaValGlyAlaThrAsnThrValPheIleLeuValAlaThrPheLeu 365
 QY 1113 ATCAGCGGCTCGGCGCGCGCTGTGTGTCGCGACGACGGCGGGGATATCTCTCC 1172
 Db 366 LeuAspArgIleArgArgProLeuValLeuThrSerThrGlyGlyMetLeuValSer 385
 QY 1173 CTCATCGGCTCGGCGCGCGGTCTACCGTCTCGCGCGACACCCCGACCAAGATACCT 1232


```

Db 386 LeuValGlyLeuAlaThrGlyLeuThrValIleSerArgHisProAspGluLysIleThr 405
QY 1233 TGGGCAATCGCCCTAAGCATCGCTCCACCTCGCTAGCTGGCTTCTTCTCATCGGC 1292
Db 406 TrpAlaIleValLeuCysIlePheCysIleMetAlaIleValAlaPhePheSerIleGly 425
QY 1293 CTGGCCCATCATCGTGGGTGTACAGTCGGAGATCTTCCGCTCCAGGTCCGCGCTG 1352
Db 426 LeuGlyProIleThrTpValIleSerSerGluIlePheProLeuHisValArgAlaLeu 445
QY 1353 GGCTGCTCGTCCGGCTCGCCGCAAGCGGTACACCGGTCCAGCGGCTCATCTCATGCTTC 1412
Db 446 GlyCysSerLeuGlyValAlaValAsnArgLeuThrSerGlyValIleSerMetThrPhe 465
QY 1413 CTGTCGTGTCCAAAGGCATCATCAGTCGGGCGAGCTTCTTCTTACTCCGCGCATCGCC 1472
Db 466 IleSerLeuSerLysAlaMetThrIleGlyAlaPhePheLeuPheAlaGlyIleAla 485
QY 1473 CGCTCGCTGGGTGTTCTTCTACACCTACTCTCCGAGACCGCGCGCGAGCTGGAG 1532
Db 486 SerPheAlaTrpValPhePheAlaIlePheAlaIlePheAlaIlePheAlaIle 505
QY 1533 GAGATGAGCAAGCTGTTCCGGCGACACGCGC-----GCCGCTCGGAATCAGAC 1580
Db 506 AspMetSerSerLeuPheGlyAsnThrAlaThrHisLysGlnGlyAlaAlaGluAlaAsp 525
QY 1581 GAGCCAGCCAAAGGAGAGAGAAAGTGGAAATGCCGCCACTAAC 1625
Db 526 AspAspAlaGlyGlu---LysLysValGluMetAlaAlaThrAsn 539

RESULT 4
ABU08338
ID ABU08338 standard; protein; 539 AA.
XX
AC ABU08338;
XX
DT 29-MAY-2003 (first entry)
XX
DE Wheat sugar transport protein #4.
XX
KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.
XX
OS Triticum aestivum.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-00051902.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
PA (ALLE// ALLEN S M.
PA (HITZ// HITZ W D.
PA (KINNEY// KINNEY A J.
PA (TING// TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR N-PSDB; ABX93210.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX
PS Claim 10; Fig 2; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis thaliana-
```

```

CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABU08334-
CC ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ Sequence 539 AA;
Alignment Scores:
Pred. No.: 9,75e-144 Length: 539
Score: 2015.00 Matches: 408
Percent Similarity: 84.4% Conservative: 44
Best Local Similarity: 76.2% Mismatches: 73
Query Match: 55.7% Indels: 10
DB: 6 Gaps: 4
US-10-051-902A-21 (1-2017) x ABU08338 (1-539)
QY 48 ACTGTACACGGCCAGAGCGA-----GCTCTCTCTCTCTGCAACCGGAG 95
Db 6 ThrGlnGlyGlyProArgArgHisAsnProGlnAlaIleAsnArgGlyAlaSerThr 25
QY 96 ATGGCTTCGCGCGCTGCGGAGCGCTGCGCGAGCGGCGGCGAGAGGCAACGTCGGTTC 155
Db 26 MetAspArgAlaAlaLeuProAlaAlaValGluProLysLysLysGlyAsnValArgPhe 45
QY 156 GCCTTCGCGCTGCGCATCTCGCTCCATGACCTCCATCCCTCGCTGCTGCTGCTGCTGCTG 215
Db 46 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 65
QY 216 GTGATGAGCGGGCGCTGCTGTATCAAGAGGACTTCAACATCAGTACGCGGAGGTG 275
Db 66 ValMetSerGlyAlaSerLeuTyrIleGlnLysAspLeuLysIleAsnAspThrGlnLeu 85
QY 276 GAGGTTCTCATGGGCATGACTGACCTCTACTCGCTCGCTCGCTCGCTCGCGCGGCGG 335
Db 86 GluValLeuMetGlyIleLeuAsnValTyrSerLeuIleGlySerPheAlaAlaGlyArg 105
QY 336 AGCTCGGACTGATCGCGCGGTACACCATCGTGTTCGCGCGCTCATATTCCTTCGCG 395
Db 106 ThrSerAspTrpIleGlyArgArgPheThrIleValPheAlaAlaValIlePheAla 125
QY 396 GGGSGGTTCTCATGGGTTCGCGTCAACTACGCTGCTCATGCTGCTGCTGCTGCTGCTG 455
Db 126 GlyAlaLeuIleMetGlyPheSerValAsnTyrAlaMetLeuMetPheGlyArgPheVal 145
QY 456 GCCGCGCATCGCGGTGCTAGCGCTCATGTCGCGCGGTGTACACCGCGAGGTGCG 515
Db 146 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValAsnThrGlyGluValSer 165
QY 516 CCGGCGTCGCGCGGTGCTTCTGACGTCTGCTCCGAGGTGTTCATCACTTCGCGATC 575
Db 166 ProAlaSerAlaArgGlyValLeuThrSerPheProGluValPheIleAsnPheGlyIle 185
QY 576 CTGCTCGGGTACGCTCTCGAATATGCTTCTCCCGCTTCCGCTGCAACTCGGCTGGCG 635
Db 186 LeuLeuGlyTyrValSerAsnPheAlaPheAlaArgLeuSerLeuArgLeuGlyTrpArg 205
QY 636 ATCATGCTCGGCATCGCGCGCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 695
Db 206 IleMetLeuGlyIleGlyValAlaValProSerValLeuLeuAlaPheMetValLeuGlyMet 225
QY 696 CCGGAGTCGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
Db 226 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 245
QY 756 GAGAGACCTCGACACGCGGAGGCGCGGAGCGCTGCGCGGAGCATCAAGCCGCC 815
```

```

Db 246 AlAlysThrSerAspThrProGluAlaAlaGluArgIleAlaAspIleLysThrAla 265
QY 816 CGCGCATCCCTGAGGAGTCGAGCGGACGTGTGACCGTCCCAAG---AGAGGGAGC 872
Db 266 AlaGlyIleProLeuGlyLeuAspGlyAspValProValProLysAsnLysGlySer 285
QY 873 GGAACAGAGAGCGGGTGTGAAGGAGCTCATCTGTCCTCCCGACCCCGGCATGCGCGC 932
Db 286 SerGluGluLysArgValLeuLysAspLeuIleLeuSerProThrIleAlaMetArgHis 305
QY 933 ATCTGTGTCTGGGATCGGATCATCTTCTTCAGCATGGTGGGATCTACTCGTC 992
Db 306 IleLeuIleAlaGlyIleGlyIlePhePheGlnGlnSerSerGlyIleAspAlaVal 325
QY 993 GTCTTCTACAGCCCTCTGCTTCAAGAGCCCGGATTAAAGACGACAAACACTTCTTG 1052
Db 326 ValLeuTyrSerProLeuValPheLysSerAlaGlyIleThrGlyAspSerArgLeuArg 345
QY 1053 GGCACCACTTGGCGCTTCGGTGTCCCAAGAGCGCTTTCATCTTGTGGGACTTCTTC 1112
Db 346 GlyThrThrValAlaValGlyAlaThrAsnThrValPheIleLeuValAlaThrPheLeu 365
QY 1113 ATGAGCGGCTCGGCGCGCGCTGTGCTGGGACGACGCGCGGATATCTCTCC 1172
Db 366 LeuAspArgIleArgArgProLeuValLeuThrSerThrGlyGlyMetLeuValSer 385
QY 1173 CTCATCGGCTCGGCGCGCGCTCACCGTCTGTCGCGCAGCACCCCGGACCAAGATACCT 1232
Db 386 LeuValGlyLeuAlaThrGlyLeuThrValIleSerArgHisProAspGluLysIleThr 405
QY 1233 TGGGCATCGGCTAAGCATCGCTCCACCTCGCTACGTCGCTTCTCTCCATCGC 1292
Db 406 TrpAlaIleValLeuCysIlePheCysIleMetAlaTyrValAlaPhePheSerIleGly 425
QY 1293 CTGTGCGGCTCAGCGGCTGACAGTCGGAGATCTTCCGCTCCAGGTGCGCGCTG 1352
Db 426 LeuGlyProIleThrTyrValTyrSerSerGluIlePheProLeuHisValArgAlaLeu 445
QY 1353 GGTCTCTCGTGGGTCGCGCGCAACCGGTCTACGCGCGCTCATCTCCATGACCTTC 1412
Db 446 GlyCysSerLeuGlyValAlaValAsnArgLeuThrSerGlyValIleSerMetThrPhe 465
QY 1413 CTGTGCGTGTCAAGCGCATCACATCGCGGCGAGTCTTCTCTCTACTCGGATCGCC 1472
Db 466 IleSerLeuSerLysAlaMetThrIleGlyIleAlaPhePheLeuPheAlaGlyIleAla 495
QY 1473 GCGTCTGCGCTGGGTGTTCTTCTACACCTTACTCCCGGAGACCCCGCGGAGCCTGGAG 1532
Db 486 SerPheAlaTyrValPhePhePheAlaTyrLeuProGluThrArgGlyArgThrLeuGlu 505
QY 1533 GAGATGACAGACTGTTCGGCGACACGCGC-----GCCGCTCGGATCAGAC 1580
Db 506 AspMetSerSerLeuPheGlyAsnThrAlaThrHisLysGlnGlyAlaAlaGluAlaAsp 525
QY 1581 GAGCAGCAGCAAGGAGAAAGAGTGAAATGGCCCGCCACTAAC 1625
Db 526 AspAspAlaGlyGlu---LysLysValGluMetAlaAlaThrAsn 539

RESULT 5
ID AAU97214 standard; protein; 529 AA.
XX AC AAU97214;
XX DT 27-AUG-2002 (first entry)
XX DE Wheat sugar transport protein encoded by wlm1.pk0012.h1.
XX KW Wheat; Beta vulgaris-like sugar transport protein;
XX KW carbohydrate transport; grain filling; annual field crop; plant.
XX OS Triticum aestivum.
XX

```

```

PN US6383776-B1.
XX 07-MAY-2002.
XX 14-APR-1999; 99US-00291922.
XX 24-APR-1998; 98US-0083044P.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2002-453364/48.
XX N-PSDB; ABK51375.
XX New nucleic acid encoding plant sugar-transport proteins, useful for
XX preparing transgenic plants with altered carbohydrate distribution.
XX Example 4; Col 77-80; 54pp; English.
XX The present invention relates to the isolation of plant polynucleotide
XX sequences encoding an Arabidopsis thaliana-like sugar transport protein
XX or Beta vulgaris-like sugar transport protein. The polynucleotide
XX sequences are useful for altering the level of sugar transport protein
XX in plants, i.e. for control of carbohydrate transport and distribution in
XX plant cells, e.g. during grain filling of annual field crops (e.g. corn,
XX rice, soybeans, and wheat), and, for studying carbohydrate flows and
XX sugar transport. The polynucleotide sequences can also be used to isolate
XX cDNA sequences and genes that encode homologues of the new proteins. The
XX present sequence represents a wheat Beta vulgaris-like sugar transport
XX protein
XX Sequence 529 AA;

Alignment Scores:
Pred. No.: 2,5e-133 Length: 529
Score: 1877.50 Matches: 369
Percent Similarity: 81.38% Conservative: 55
Best Local Similarity: 70.83% Mismatches: 92
Query Match: 51.95% Indels: 5
DB: Gaps: 3

US-10-051-902A-21 (1-2017) x AAU97214 (1-529)
QY 60 CCAGAGCGA-----GCTCTCTCTCTGACACCGGAGATGCTTCCGCCGCTG 113
Db 5 ProGluArgLysGlyAlaGluAspLysGluGluGlySerArgMetAlaSerAlaAlaLeu 24
QY 114 CCGGAG-----GCCGTGCGCGCGGAGAGGAGGCAACGTCGCTTCCCTCCCTGC 167
Db 25 ProGluProGlyAlaValHisProArgAsnLysGlyAsnPheLysTyrAlaPheThrCys 44
QY 168 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 227
Db 45 AlaLeuCysAlaSerMetAlaThrIleValLeuGlyTyrAspValGlyValMetSerGly 64
QY 228 GCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 287
Db 65 AlaSerLeuTyrIleLysArgAspLeuGluIleThrAspValGlnLeuGluIleMetMet 84
QY 288 GGCATCTGCACTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 347
Db 85 GlyIleLeuSerValTyrAlaLeuIleGlySerPheLeuGlyAlaArgThrSerAspTrp 104
QY 348 ATCGCGCGCGGTACACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407
Db 105 ValGlyArgArgValThrValPheAlaAlaAlaIlePheAsnAsnGlySerLeuLeu 124
QY 408 ATGGGGTTCGCGCTCACTACCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
Db 125 MetGlyPheAlaValAsnTyrAlaMetLeuMetValGlyArgPheValThrGlyIleGly 144
QY 468 GTGGGCTACGCTCATGATCGCGCGGTGTACACCGCGAGGTGTCGCGCGGTGCGCG 527

```

Db	145	ValGlyTyrAlaIleMetValAlaProValTyrThrProGluValSerProAlaSerAla	164
Qy	528	CGTGGCTTCCTGACGTCTCCGAGAGGTGTTCAATCAACTCGGCATCTCTCGGGTAC	587
Db	165	ArgGlyPheLeuThrSerPheThrGluValPheIleAsnValGlyIleLeuLeuGlyTyr	184
Qy	588	GTCTCGAACAATGTTTTCTCCGCTTCGCGTGAACCTCGGGTGGCGCATCATGCTCGG	647
Db	185	ValSerAsnTyrAlaPheAlaArgLeuProLeuHisLeuSerTrpArgValMetLeuGly	204
Qy	648	ATCGGCGCGCGCGCTCGCTGCTCGCGTCAATGCTGCTCGGCATCGCGGAGTCGCG	707
Db	205	IleGlyAlaValProSerAlaLeuLeuAlaLeuMetValPheGlyMetProGluSerPro	224
Qy	708	CGTGTGCTGTCATGAAGGACGCTCCGCGAGCCCAAGTGTGTGGAGAAAGACCTCC	767
Db	225	ArgTrpLeuValMetLysGlyArgLeuAlaAspAlaArgAlaValLeuAlaLysThrSer	244
Qy	768	GACACGCGGAGGCGCGGCGCGCTCGCGCACATCAAGCGCGCGCGGATCCCT	827
Db	245	AspTrpProGluGluAlaValGluArgLeuAspGlnIleLysAlaAlaGlyIlePro	264
Qy	828	GAGGAGCTCAGCGCGACGTGGTCAACGCTCCCAAGAGAGGAGCGGAAACGAGAACGG	887
Db	265	ArgGluLeuAspGlyAspValValMetProLysThrLysGlyGlyGlnLysGln	284
Qy	888	GTGTGGAAGAGCTCATCTGTCGCCACCCCGCCATGGCGGCATCTCTGTCGCGG	947
Db	285	ValTrpLysGluLeuIlePheSerProAlaMetArgArgIleLeuLeuAlaAla	304
Qy	948	ATCGGCATCCACTTCTCCACGATCGCTGGGCATTCACTCCGTCGCTCTCTACAGCCCT	1007
Db	305	LeuGlyIleHisPhePheGlnAlaThrGlySerAspSerValValLeuTyrSerPro	324
Qy	1008	CTCGTGTTCAGAGCCCCGGATTAACGAACGAAACACTTCTTGGGACACACTTGGCGG	1067
Db	325	ArgValPheGlnSerAlaGlyIleThrGlyAspAsnHisLeuLeuGlyAlaThrCysAla	344
Qy	1068	TTGCGTGTACCAAGAGCTTTTCATCTGTGTGGCGACTTCTTTCATCGAGCGGCTCGG	1127
Db	345	MetGlyValMetLysThrLeuPheIleLeuValAlaThrPheGlnLeuAspArgValGly	364
Qy	1128	CGGCGCGCGTGTGTGGGACGACCGGGGGGATAATCTCTCCCTCATCTCGGCCTCGGC	1187
Db	365	ArgArgProLeuLeuLeuThrSerThrAlaGlyMetLeuAlaCysLeuIleGlyLeuGly	384
Qy	1188	GCGCGGCTCACGTCGTCGCGCACGCCCGACCGCAGATACCTTGGCCATCGGCCTA	1247
Db	385	ThrGlyLeuThrValValGlyArgHisProAspAlaLysValProTrpAlaIleGlyLeu	404
Qy	1248	AGCATCGCCTCCACCCCTAGCTCGCCTCTTCTCCATCGGCCTTGGCCCATCAGC	1307
Db	405	CysIleValSerIleLeuAlaTyrValSerPhePheSerIleGlyLeuGlyProLeuThr	424
Qy	1308	TGGGTGTACAGCTCGGAGATCTTCCGCTCCAGTGGCGCGCTGGGCTCTCGCTCGGC	1367
Db	425	SerValTyrThrSerGluValPheProLeuArgValArgAlaLeuGlyPheAlaLeuGly	444
Qy	1368	GTCGCGCCAAACCGGCTCACGAGCGGCGTCAATCCATGACCTTCTCTGTCTGCTCCAAG	1427
Db	445	ThrSerCysAsnArgValThrSerAlaAlaValSerMetSerPheLeuSerLeuSerLys	464
Qy	1428	GCCATACCAATCGGGGAGCTTCTTCTCTACTCTCGGCATCGCGCGCTCGCTCGGGTG	1487
Db	465	AlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyIleAlaAlaIleGlyTrpIle	484
Qy	1488	TTCTTCTACACTACCTCCGAGACCGCGCGCGGACGCGCTGGAGGAGATGAGCAAGCTG	1547
Db	485	PhePhePheThrPheIleProGluThrArgGlyLeuProLeuGluGluIleGlyLysLeu	504
Qy	1548	TTGCGGCACACGCGCGCGCTCGGAATCAGACGAGCAGCC---AAGAGGAAGAAGAAG	1604

Db 505 PheGlyMetThrAspThrAlaValGluAlaGlnAspThrAlaThrIysAspLysAlaLys 152

Qy 1605 GTG 1607
|||

Db 525 Val 525

RESULT 6

ABU08339

ID ABU08339 standard; protein; 529 AA.

XX

AC ABU08339;

DT 29-MAY-2003 (first entry)

XX

DE Wheat sugar transport protein #5.

XX

KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
plant sugar transport protein; carbohydrate transport; soybean;
carbohydrate distribution; plant.

XX

OS Triticum aestivum.

OS

PN US2002178468-A1.

XX

XX

PD 28-NOV-2002.

XX

PF 17-JAN-2002; 2002US-00051902.

XX

PR 24-APR-1998; 98US-0083044P.

PR 14-APR-1999; 99US-00291922.

XX

PA (ALLE/) ALLEN S M.
(HITZ/) HITZ W D.
(KINN/) KINNEY A J.
(TINGE/) TINGEY S V.

XX

PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

XX

DR WPI; 2003-340957/32.

DR N-PSDB; ABX93211.

XX

PT Novel plant sugar transport proteins and nucleic acid encoding the
protein useful for producing transgenic plants having altered levels of
sugar transport protein.

XX

PS Claim 10; Fig 2; 56pp; English.

XX

CC The present invention relates to the isolation of Arabidopsis thaliana-
like or Beta vulgaris-like sugar transport proteins, and the
polynucleotide sequences encoding them. The plant sugar transport
proteins of the invention have been isolated from corn, rice, soybean, the
and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
transforming a host cell with a chimeric construct encoding all, or a
portion of the sugar transport protein, in sense or antisense
orientation. Particularly, the polypeptides may provide a means to
control carbohydrate transport and distribution in plants. ABU08334-
ABU08339 represent Beta vulgaris-like sugar transport proteins

XX

SQ Sequence 529 AA;

Alignment Scores:

Pred. No.: 2,5e-133 Length: 529

Score: 1877.50 Matches: 369

Percent Similarity: 81.38% Conservative: 55

Best Local Similarity: 70.83% Mismatches: 92

Query Match: 51.95% Indels: 5

DB: Gaps: 3

US-10-051-902A-21 (1-2017) x ABU08339 (1-529)

Qy 60 CCGAGACGA-----GCCCTCTCTCTCTGACACCGAGATGGCTTCGGCGGCGCTG 153
|||||

Db 5 ProGluAArgLysGlyAlaGluAspLysGluGluGlySerArgMetAlaSerAlaAlaLeu 24
 QY 114 CCGGAG-----GCCGTCCGCGCGAAGAAGAGGCAACGTCCTCGCTTCGCTCGCCTGC 167
 Db 25 ProGluProGlyAlaValHisProArgAsnLysGlyAsnPheLysTyrAlaPheThrCys 44
 QY 168 GCCATCCCTCGCTCATACCTCCATCTCTCTCGCTACGATATCGGGTGATAGCGGG 227
 Db 45 AlaLeuCysAlaSerMetAlaThrIleValLeuGlyTyrAspValGlyValMetSerGly 64
 QY 228 GGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGAGGAGTGAGTCTCATG 287
 Db 65 AlaSerLeuTyrIleLysArgAspLeuGlnIleThrAspValGlnLeuGluIleMetMet 84
 QY 288 GGCATACCTGAACCTACTCGCTCATCGCTCTCTCGCTCGCGCGCGGCGGACGTCGACTGG 347
 Db 85 GlyIleLeuSerValTyrAlaLeuIleGlySerPheLeuGlyAlaArgThrSerAspTrp 104
 QY 348 ATCGCGCGCGGTACACCATCGTGTTCGCGCGCGCTCATATTCTTCGCGGGSGGTTCCTC 407
 Db 105 ValGlyAArgValThrValPheAlaIleAlaIlePheAsnAsnGlySerLeuLeu 124
 QY 408 ATGGGTTTCGCGCTCAACTACGCCATGCTCATGTTTCGCGCGCTTCGTCGGCGGCGATCGCG 467
 Db 125 MetGlyPheAlaValAsnTyrAlaMetLeuMetValGlyArgPheValThrGlyIleGly 144
 QY 468 GTGGCTACGGCTCATGATCGCGCGGTGTACACCGCGGAGTGTCCGCGGTTCGCGG 527
 Db 145 ValGlyTyrAlaIleMetValAlaProValTyrThrProGluValSerProAlaSerAla 164
 QY 528 CQTGCTTCCTCGACGCTCTCCGAGGTGTTCATCACTTCGGCATCCTCGCTCGGCTAC 587
 Db 165 ArgGlyPheLeuThrSerPheThrGluValPheIleAsnValGlyIleLeuLeuGlyTyr 184
 QY 588 GTCTCGAATGCTTTCTCCGCTTCGCTGCTGAACTCGGCTCGGCTCGGCTCGGCTCGG 647
 Db 185 ValSerAsnTyrAlaPheAlaArgLeuProLeuHisLeuSerTrpArgValMetLeuGly 204
 QY 648 ATCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707
 Db 205 IleGlyAlaValProSerAlaLeuLeuLeuLeuMetValPheGlyMetProGluSerPro 224
 QY 708 CGTGGCTGTGATGAAGGACGCTCGCGACGCCAAGTGTGCTGCGAGAGACCTCC 767
 Db 225 ArgTrpLeuValMetLysGlyArgLeuAlaAspAlaArgAlaValLeuAlaLysThrSer 244
 QY 768 GACAGCGGAGGAGGCGCGGAGCGCTGCGGACATCATGCGCGCGCGCGCGCGCATCCCT 827
 Db 245 AspThrProGluGluAlaValGluArgLeuAspGlnIleLysAlaAlaGlyIlePro 264
 QY 828 GAGGAGCTCGACGCGGCTGTGACCGCTCCCAAGAGAGGAGCGGAGGAGGAGGAGGAGG 887
 Db 265 ArgGluLeuAspGlyAspValValMetProLysThrLysGlyGlyGlnGluLysGln 284
 QY 888 GTGGAGAGGAGCTCATCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 947
 Db 285 ValTrpLysGluLeuIlePheSerProThrProAlaMetArgArgIleLeuLeuAlaAla 304
 QY 948 ATCGGCATCCACTTCTCCAGCATGCGTGGGCACTTCATCCGCTGCTTCTACAGCCCT 1007
 Db 305 LeuGlyIleHisPhePheGlnAlaThrGlySerAspSerValValLeuTyrSerPro 324
 QY 1008 CTCGTGTTCAAGAGCCCGGATTAACGAGACGAGAAACACTTCTTGGGCGACCATCTGGCGG 1067
 Db 325 ArgValPheGlnSerAlaGlyIleThrGlyAspAsnHisLeuLeuGlyAlaThrCysAla 344
 QY 1068 TTCGTGTGCACAGAGCTTTTCATCTGTGTGCGACTTCTTCTCATCGCGCGGCTCGGG 1127
 Db 345 MetGlyValMetLysThrLeuPheIleLeuValAlaThrPheGlnLeuAspArgValGly 364
 QY 1128 CGGCGCGCGCTGTGCTGGGAGCAGCGCGGAGGATATCTCTCTCCCTCATCGGCTCGG 1187
 Db 365 ArgArgProLeuLeuThrSerThrAlaGlyMetLeuAlaCysLeuIleGlyLeuGly 384

QY 1188 GCCGGGCTCACCGTCTCGGCGAGCACCCGACGACGACGACGACGACGACGACGACGACG 1247
 Db 385 ThrGlyLeuThrValValGlyArgHisProAspAlaLysValProTrpAlaIleGlyLeu 404
 QY 1248 AGCATCGCTCCACCCCTCGCTACGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1307
 Db 405 CysIleValSerIleLeuAlaIleValSerPhePheSerIleGlyLeuGlyProLeuThr 424
 QY 1308 TGGGTGTACAGCTCGGAGATCTCCGCTCCAGGTCCGCGCGCGCTCGCTCGCTCGCTCG 1367
 Db 425 SerValTyrThrSerGluValPheProLeuArgValArgAlaLeuGlyPheAlaLeuGly 444
 QY 1368 GTGCGCGCGCAACCGGTACCGCGCGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1427
 Db 445 ThrSerCysAsnArgValThrSerAlaAlaValSerMetSerPheLeuSerLeuSerLys 464
 QY 1428 GCATCACCATCGGCGGCGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1487
 Db 465 AlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyIleAlaIleGlyTrpIle 484
 QY 1488 TTCTTCTACACCTACTCTCCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1547
 Db 485 PhePhePheThrPheIleProGluThrArgGlyLeuProLeuGluGluIleGlyLysLeu 504
 QY 1548 TTCGGCGACACGCGCGCGCTCGGATCAGACGAGCGCGCGCGCGCGCGCGCGCGCGCG 1604
 Db 505 PheGlyMetThrAspThrAlaValGluAlaGlnAspThrAlaThrLysAspLysAlaLys 524
 QY 1605 GTG 1607
 Db 525 Val 525
 RESULT 7
 AAU97210
 ID AAU97210 standard; protein; 513 AA.
 XX AC AAU97210;
 XX 27-AUG-2002 (first entry)
 XX XX
 DE Corn Beta vulgaris-like sugar transport protein.
 XX XX
 KW Corn; Beta vulgaris-like sugar transport protein; carbohydrate transport;
 XX grain filling; annual field crop; plant.
 XX Zea mays..
 OS US6393776-Bi.
 FN 07-MAY-2002.
 XX 14-APR-1999; 99US-00291922.
 XX 24-APR-1998; 98US-0083044P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX WPI: 2002-453364/48.
 DR N-PSDB; ABX51371.
 XX
 PT New nucleic acid encoding plant sugar-transport proteins, useful for
 XX preparing transgenic plants with altered carbohydrate distribution.
 XX Example 4; Fig 2; 54pp; English.
 PS The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
 CC or Beta vulgaris-like sugar transport protein. The polynucleotide
 CC sequences are useful for altering the level of sugar transport proteins
 CC in plants, i.e. for control of carbohydrate transport and distribution in

PR 14-APR-1999; 99US-00291922.
XX (ALLEN/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINNE/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2003-340957/32.
DR N-PSDB; ABX93207.
XX
XX Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX
XX Claim 10; Fig 2; 56pp; English.
XX
XX The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABU08334-
CC ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ Sequence 513 AA;

Alignment Scores:
Pred. No.: 5.9e-133 Length: 513
Score: 1872.50 Matches: 372
Percent Similarity: 84.38% Conservative: 60
Best Local Similarity: 72.66% Mismatches: 77
Query Match: 51.81% Indels: 3
DB: 6 Gaps: 2

US-10-051-902A-21 (1-2017) x ABU08335 (1-513)
QY 96 ATGGCTTCGGCGGCTCCGCGAGCGCTCCGCGCGAAGAGGCAACGTCGCGTTC 155
DB 1 MetAlaSerAspGluLeuAlaAlaMetTyrIleLysAspValLysTyr 20
QY 156 GCCTTCGCTCGGATCCTCGCTCCATGACCTCCATCTCTCTCGGCTAGCATCGGG 215
DB 21 AlaSerIleCysAlaIleLeuAlaSerMetAlaSerValIleLeuGlyTyrAspIleGly 40
QY 216 GTGATGAGCGGGGCTGCTGTACATCAAGAGACTTCAACATCAGTACGGGAAGGTG 275
DB 41 ValMetSerGlyAlaAlaMetTyrIleLysAspLeuAsnIleThrAspValGlnLeu 60
QY 276 GAGGTTCATGGGCATCTCAACTCACTGCTCATGCTCATGCTTCGCGGGGGCGG 335
DB 61 GluIleLeuIleGlyIleLeuSerLeuTyrSerLeuPheGlySerPheAlaGlyAlaArg 80
QY 336 ACCTCGGACTGATCGGCGGCTGATCACCATGCTGTTCGCGCGTCATATCTTCGCG 395
DB 81 ThrSerAspArgIleGlyArgLeuThrValPheAlaAlaValIlePhePheVal 100
QY 396 GGGSGTTCATCGGGGTTCGCGTCACTACGCTCATGCTCATGCTTCGCGCGCTTCGTG 455
DB 101 GlySerLeuLeuMetGlyPheAlaValAsnTyrGlyMetLeuMetAlaGlyArgPheVal 120
QY 456 GCCGCGATCGGCGGCTAGCGCTCATGATCGCGCGGTGTACACCGCGGAGGTGTGCG 515
DB 121 AlaGlyValGlyValGlyTyrGlyGlyMetIleAlaProValTyrThrAlaGluIleSer 140
QY 516 CCGCGCTCGGCGGCTGCTCTGACGTCTGCTCCGCGAGGTGTTCATCACTTCGCGATC 575
DB 141 ProAlaAlaSerArgGlyPheLeuThrThrPheProGluValPheIleAsnIleGlyIle 160

QY 576 CTGCTCGGCTAGCTCTCGAACTATGCTTTCTCCGCTTCCGCTGAACCTCGGTCGGCGC 635
DB 161 LeuLeuGlyTyrLeuSerAsnPheAlaPheAlaArgLeuProLeuHisLeuGlyTyrArg 180
QY 636 ATCATGCTCGGATCGGCGGCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 695
DB 181 ValMetLeuAlaIleGlyAlaValProSerGlyLeuLeuAlaLeuValPheCysMet 200
QY 696 CCGGAGTCCGCGGTGGCTGTATGAAGGAGCGCTCGCGAGCGCCAAAGCTGGTGTG 755
DB 201 ProGluSerProArgTyrLeuValLeuLysGlyArgLeuAlaAspAlaValLeu 220
QY 756 GAGAAGACTCCGACACGCGGAGGAGCGCGGAGCGCTCGCGACATCAAGCGCGCC 815
DB 221 GluLysThrSerAlaThrProGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240
QY 816 GCGGCTATCCCTGAGGAGCTCGACGCGAGCTGTCGCTCCGCTCCGCTCCGCTCCG 872
DB 241 AlaGlyIleProLysGlyLeuAspGlyAspValValThrValProGlyLysGluGlnGly 260
QY 873 GGAACGAGAGCGGCTGTGGAAGAGCTCATCTCTCCGACCGCGCCATCGCGCGC 932
DB 261 GlyGlyGluLeuGlnValTyrLysLeuIleLeuSerProThrProAlaValArgArg 280
QY 933 ATCTGCTGCTCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATC 992
DB 281 IleLeuLeuSerAlaValGlyLeuHisPhePheGlnGlnAlaSerGlySerAspSerVal 300
QY 993 GTCTTCTACAGCCCTCTCGTGTCAAGAGCGCGGATTAACGAAACGAAACACATCTCT 1052
DB 301 ValGlnTyrSerAlaArgLeuPheLysSerAlaGlyIleThrAspAspAsnLysLeuLeu 320
QY 1053 GGCACCATCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 1112
DB 321 GlyValThrCysAlaValGlyValThrLysThrPhePheIleLeuValAlaThrPheLeu 340
QY 1113 ATCGACGGCTCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1172
DB 341 LeuAspArgAlaGlyArgProLeuLeuLeuIleSerThrGlyGlyIleValSer 360
QY 1173 CTCTATGCGCTCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1232
DB 361 LeuIleCysLeuGlySerGlyLeuThrValAlaGlyHisHisProAspThrLysValAla 380
QY 1233 TGGGCTATCGGCTAAGCATCGCTCCACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1292
DB 381 TrpAlaValAlaLeuCysIleAlaSerThrLeuSerTyrIleAlaPheSerIleGly 400
QY 1293 CTTGCGCCCATCACCTGGGTGTACAGCTCGGAGATCTTCCGCTCCAGTCCGCGCTG 1352
DB 401 LeuGlyProIleThrGlyValTyrThrSerGluIlePheProLeuGlnValArgAlaLeu 420
QY 1353 GCTCTGCTGCTGCGGCTGCGCGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1412
DB 421 GlyPheAlaValGlyValAlaSerAsnArgValThrSerAlaValIleSerMetThrPhe 440
QY 1413 CTGCTGCTGTCCAAGCGCATCACCATCGGCGGAGCTTCTCTCTCTCTCTCTCTCTCT 1472
DB 441 LeuSerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAla 460
QY 1473 GCGCTGCTGCTGCTGCTTCTTCTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1532
DB 461 AlaValAlaIleValPhePhePheThrCysLeuProGluThrArgGlyArgThrLeuGlu 480
QY 1533 GAGATGAGCAAGCTGCTCGGC-----GACACGCGCGCGCTCGGATCAGACGAGCA 1586
DB 481 GluMetGlyLysLeuPheGlyMetProAspThrGlyMetAlaGluGluAlaGluAspAla 500
QY 1587 GCCAAGGAG 1622
DB 501 AlaAlaLysGluLysValValGluLeuProSerSer 512

```

RESULT 9
AAU97212
ID AAU97212 standard; protein; 523 AA.
XX AC
XX AAU97212;
XX
DT 27-AUG-2002 (first entry)
XX
DE Soybean Beta vulgaris-like sugar transport protein.
XX
KW Soybean; Beta vulgaris-like sugar transport protein;
XX carbohydrate transport; grain filling; annual field crop; plant.
XX
OS Glycine max.
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
FF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2002-453364/48.
XX
DR N-PSDB; ABKS1973.
XX
XX
XX New nucleic acid encoding plant sugar-transport proteins, useful for
XX preparing transgenic plants with altered carbohydrate distribution.
XX
XX Example 4; Fig 2; 54pp; English.
XX
XX The present invention relates to the isolation of plant polynucleotide
XX sequences encoding an Arabidopsis thaliana-like sugar transport protein
XX or Beta vulgaris-like sugar transport protein. The polynucleotide
XX sequences are useful for altering the level of sugar transport proteins
XX in plants, i.e. for control of carbohydrate transport and distribution in
XX plant cells, e.g. during grain filling of annual field crops (e.g. corn,
XX rice, soybeans, and wheat), and, for studying carbohydrate flows and
XX sugar transport. The polynucleotide sequences can also be used to isolate
XX cDNA sequences and genes that encode homologues of the new proteins. The
XX present sequence represents a soybean Beta vulgaris-like sugar transport
XX protein
XX
XX SQ Sequence 523 AA;
XX
Alignment Scores:
Pred. No.: 7.31e-117 Length: 523
Score: 1660.00 Matches: 331
Percent Similarity: 75.76% Conservative: 69
Best Local Similarity: 62.69% Mismatches: 110
Query Match: 45.93% Indels: 18
DB: 5 Gaps: 4
XX
US-10-051-902A-21 (1-2017) x AAU97212 (1-523)
QY 99 GCTTCGCGCGCTGCGGAGCGCTGCGCGGAGAGAGGCGCAACGTCGGTTCGCC 158
Dz 12 AlahLysThrLeuGlnAspPheAspProProLysLysArgLysArgAsnLysTyrAla 31
QY 159 TTCGCTCGCGCATCCTCGCTCCATGACCTCCATCCTCTCGGCTACGATATCGGGTG 218
Dz 32 PheAlaCysAlaMetLeuAlaSerMetThrSerileLeuLeuGlyTyrAspIleGlyVal 51
QY 219 ATGAGCGGGGCTGCTGTATCATCAAGAGACTTCAACATCACTGACGGGAAGGTGGAG 278
Dz 52 MetSerGlyAlaAlaileTyrileLysArgAspLeuLysValSerAspGluGlnileGlu 71
QY 279 GTTCTCATGGCATACTGAACCTTACTCGCTCATCGCTCTTTCGCGGCGGGCGGACG 338

```

```

Db 72 IleLeuLeuGlyIleIleAsnLeuTyrSerLeuIleGlySerCysLeuAlaGlyArgThr 91
QY 339 TCGGACTGGATCGCGCGCGGTACACCATCGGTTCGCGCGCGTTCATATTCCTCGCGGG 398
Dz 92 SerAspTrpIleGlyProArgTyrThrIleValPheAlaGlyThrIlePhePheValGly 111
QY 399 GSGTTCCTCATCGGGTTCGCCGTCACCTACCGCATGCTTCATGTTGGCGGCTTCGTGGCC 458
Dz 112 AlaLeuLeuMetGlyPheSerProAsnTyrSerPheLeuMetPheGlyArgPheValAla 131
QY 459 GGCATCGGCGTGGCTACGCGCTCATGATCGCGCGGTGTACACCGCGGAGGTGTGCGCG 518
Dz 132 GlyIleGlyIleGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerPro 151
QY 519 GCGTCGCGCGTGGCTTCCTGACGTCGTTCGCGGAGGTGTTCATCAACTCGGCATCTG 578
Dz 152 AlaSerSerArgGlyPheLeuThrSerPheProGluValPheIleAsnGlyIleLeu 171
QY 579 CTCGGGTACGTCGAACTATGCTTCTCCGCTTCGCGCTGAACCTCGGGTGGCGCATC 638
Dz 172 IleGlyTyrIleSerAsnTyrAlaPheSerLysLeuThrLeuLysValGlyTyrPargMet 191
QY 639 ATGCTCGGCATCGCGCGCGCGCTCGGTTCGTGCTCGCGCTCATGCTGCTCGCATGCG 698
Dz 192 MetLeuGlyValGlyAlaIleProSerValLeuLeuThrValGlyValLeuAlaMetPro 211
QY 699 GAGTCGCGCGTGGCTGTCATGAAGGAGCGCTCGCGGAGCCCAAGGTGCTGTGGAG 758
Dz 212 GluSerProArgTrpLeuValMetArgGlyArgLeuGlyGluAlaArgLysValLeuAsn 231
QY 759 AAGACTCTCGACACGCGGAGGAGCGCGGAGCGCTCGCGGAGCATCAAGCGCGCGCC 818
Dz 232 LysThrSerAspSerLysGluGluAlaGlnLeuArgLeuAlaGluIleLysGlnAlaAla 251
QY 819 GGCATCCCTGAGAGCTCGACGCGGACGTGTGTGACGCTCCCGCCCAAGAGGGAGCGGAAC 878
Dz 252 GlyIleProGluSerCysAsnAspValValGlnValAlaAsnLysGlnSerAsnGly--- 270
QY 879 GAGAGCGCGTGTGAAGGAGCTCATCTCTGCCGACCCCGCGCATCGCGGCGCATCTCG 938
Dz 271 ---GluGlyValTrpLysGluLeuPheLeuTyrProThrProAlaIleArgHisIleVal 289
QY 939 CTGTCCGGATCGGATCCATCTTCCAGCATGCGTGTGGGATTCATCTCCCTGCTTC 998
Dz 290 IleAlaAlaLeuGlyIleHisPhePheGlnAlaSerGlyValAspAlaValValLeu 309
QY 999 TACAGCCCTCTCGTGTCAAGAGCCCGGATTAACGAGCAGCAACACATCTTCTGGGCACC 1058
Dz 310 TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuAla 329
QY 1059 ACTTGGCCCTTCGCTGTCACCAAGAGGCTTTTTCATCTTGTTCGCGCATTTCTTCATCGAC 1118
Dz 330 ThrValAlaValGlyPheValLysThrValPheIleLeuAlaAlaThrPheThrLeuAsp 349
QY 1119 GCGTCGCGCGCGCGCTGTGCTGGGACACGCGGGGATATCTCTCCCTCATC 1178
Dz 350 ArgValGlyArgArgProLeuLeuLeuSerValGlyGlyMetValLeuSerLeuLeu 369
QY 1179 GCGCTCGGCGCGGCTCACCGCTCGCGGACGCGCGGACCGCAAGATACCTTGGGCGC 1238
Dz 370 ThrLeuAlaIleSerLeuThrValIle---AspHisSerGluArgLysLeuMetTrpAla 388
QY 1239 ATCGGCTTAAGCATCGCTCCACCTCGCTACGCTCGCTTCTTCTCCATCGGCTTGGC 1298
Dz 389 ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly 408
QY 1299 CCGATCAGTGGGTGTACAGCTCGGAGATCTTCCGCTCCAGTCCGCGGCTGGCTGC 1358
Dz 409 ProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnIleGlu 428
QY 1359 TCGCTCGGCGTCCGCGCAACCGCTCACCGAGCGGTCTATCTCCATGACCTTCTGTG 1418
Dz 429 AlaAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448

```


QY	999	TACAGCCCTCTCGTGTTCAAGAGCCCGGATTACGAAACGACAAACACACTTCTTGGGGACC	1058
Db	310	TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuAla	329
QY	1059	ACTTGGCGCGTTCGCTGTACACAGAGCGCTTTTCATCTTGTGGCGACTTCTTCATCGAC	1118
Db	330	ThrValAlaValGlyPheValLysThrValPheIleLeuAlaIleThrPheThrLeuAsp	349
QY	1119	GGCGTCGGGGGGCGCGCTGTGCTGGGACGACGGGGGGGATATCTCTCTCCCTCATC	1178
Db	350	ArgValGlyArgArgProLeuLeuLeuSerSerValGlyGlyMetValLeuSerLeuLeu	369
QY	1179	GGCGTCGGCGCGCGCTCACCGTCGTCGGCCAGCACCCCGACCAAGATACCTTGGGCC	1238
Db	370	ThrLeuAlaIleSerLeuThrValIle--AspHisSerGluargLysLeuMetTrpAla	388
QY	1239	ATCGGCTAAGCATCGGCTCCACCTCGCTACGTCGCTTCTTCTCCATCGCGCTTGGC	1298
Db	389	ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly	408
QY	1299	CCCATCAGCGGGGTACAGCTCGGAGACTTCCCGCTCCAGGTGCGCGCTGGGCTGC	1358
Db	409	ProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnGlyAla	428
QY	1359	TCGCTCGGCTCGCGCCAAACCGCGTCACCGAGCGGCGTCACTCCATGACTTCTCTGTCG	1418
Db	429	AlaAlaGlyValAlaValAsnAsgThrThrSerAlaValValSerMetThrPheLeuSer	448
QY	1419	CTGTCCAAAGCCATCACCATCGGCGGAGCTTCTTCTCTACTCCGCGATCGCGCGCTC	1478
Db	449	LeuThrArgAlaIleThrIleGlyGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal	468
QY	1479	GCCTGGGTGTCTTCTACACTACTCTCCCGAGACCCGCGCGGAGCGCTGGAGGAGATG	1538
Db	469	GlyTrpIlePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet	488
QY	1539	AGCAAGCTGTTGGCGGACACGGCGCGCGCTCGGAATCAGACGACGACCAAGGAGAAG	1598
Db	489	GlugLysPheGlyThrPheArgSerLysSerAsnAla-----Ser	502
QY	1599	AAGAAGTGGAAATGGCGCCACTACTGATCAAACTAACCGCAAAATCACCAAATCCCTA	1658
Db	503	LysAlaValGlu-----AsnGluAsnGlyGlnValAla	513
QY	1659	AGGGTTTCTTGCAAAAACGTGTG	1682
Db	514	GlnValGlnLeuGlyThrAsnVal	521
RESULT 11			
ABB09681			
XX	ID	ABB09681	standard; protein; 513 AA.
XX	AC	ABB09681;	
XX	DT	11-JUN-2002	(first entry)
XX	DE	Amino acid sequence of a mannitol transporter designated AgMat1.	
KW	KW	Linear polyol; mannitol; mannitol transporter; carbon source; plant;	
KW	KW	pathogen resistance; salt stress; sorbitol; dulcitol; galactitol;	
KW	KW	inositol; ribitol; xylitol; AgMat1.	
XX	OS	Apium graveolens.	
XX	PN	WO200204547-Al.	
XX	PD	17-JAN-2002.	
XX	XX	22-JUN-2001; 2001WO-FR001979.	
XX	PF		
XX	PR	11-JUL-2000; 2000FR-00009032.	
XX	PA	(CNRS) CENT NAT RECH SCI.	


```
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-014091P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142940P.
PR 12-JUL-1999; 99US-0143542P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147935P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148585P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.

PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 08-OCT-1999; 99US-0157865P.
PR 08-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 4.35e-96 Length: 491
Score: 1385.50 Matches: 271
Percent Similarity: 72.41% Conservative: 78
Best Local Similarity: 56.22% Mismatches: 120
Query Match: 38.34% Indels: 13
DB: 3 Gaps: 3

US-10-051-902A-21 (1-2017) x AAG32072 (1-491)
QY 150 CGGTCGCTTCGCTGCGCCATCTCGCTCCATGACTCCATCTCCTCGGTACCAT 209
Db 3 LysPheAlaPheGlyCysAlaIleValAlaSerIleIleSerIleIlePheGlyTyrAsp 22
QY 210 ATCGGGGTGATGAGCGGGCGTGTACATCAAGAGGACTTCAACATCAGTCACGG 269
Db 23 ThrGlyValMetSerGlyAlaGlnIlePheIleArgAspAspLeuLysIleAsnAspThr 42
QY 270 AAGGTGGAGGTCTCATGGGCATCTGAACCTTACTCGCTCATCGGTCTCTCGCGCG 329
Db 43 GlnIleGluValLeuAlaGlyIleLeuAsnLeuCysAlaLeuValGlySerLeuThrAla 62
QY 330 GGGCGGACGTGGACTGGATCGGGCGGGTACACCATCGGTGTCGGCGCGCTCATATC 389
Db 63 GlyLysThrSerAspValIleGlyArgArgTyrThrIleAlaLeuSerAlaValIlePhe 82
QY 390 TTCGGCGGGGGTCTCTCATGGGTTCGCCGCTCACTACGCATGCTCATGTTGGCGCG 449
```

Db 83 LeuValGlySerValLeuMetGlyTyrGlyProAsnTyrProValLeuMetValGlyArg 102
Qy 450 TTCGTGGCGGCATCGGGTGGGTACCGCTCATGATCGCGCGGTGTACACCGCGAG 509
Db 103 CysIleAlaGlyValGlyValGlyPheAlaLeuMetIleAlaProValTyrSerAlaGlu 122
Qy 510 GTGTGCGCGGCGTGGCGGTGGTCTCTGACGTCGTTCGCGAGGTGTTCATCAACTTC 569
Db 123 IleSerAlaSerHisArgGlyPheLeuThrSerLeuProGluLeuCysIleSerLeu 142
Qy 570 GGCATCTGTCGGGTACGTCCTGAACTGCTTCTCCCGCTTCCGCTGACCTGACCTCGG 629
Db 143 GlyIleLeuLeuGlyTyrValSerAsnTyrCysPheGlyLysLeuThrLeuLysLeuGly 162
Qy 630 TGGCGCATCATCGGCGCGCGCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 689
Db 163 TrpArgLeuMetLeuGlyIleAlaAlaPheProSerLeuLeuLeuAlaPheGlyIleThr 182
Qy 690 GGCATGCGGAGTCCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
Db 183 ArgMetProGluSerProArgTrpLeuValMetGlnGlyArgLeuGluGluAlaLysLys 202
Qy 750 GTGCTGGAGAGACTCCGACACCGCGGAGGAGCGCGGAGCGCTGCGGACATCAAG 809
Db 203 IleMetValLeuValSerAsnThrGluGluAlaGluArgPheArgAspIleLeu 222
Qy 810 GCCGCGCGCGCATCCCTGAGGAGTCCGACGCGGAGCGGTGCTGCTGCTGCTGCTGCTGCT 869
Db 223 ThrAlaAlaGluVal-----AspValThrGluIleLysGluValGly 236
Qy 870 ACCGGAACAGAGACCGG-----GTGTGGAAGGAGCTCATCTGCTCCCGC 914
Db 237 GlyGlyValLysLysLysAsnHisGlyLysSerValTrpArgGluLeuValIleLysPro 256
Qy 915 ACCCGCGCATCGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 974
Db 257 ArgProAlaValArgLeuLeuIleAlaValGlyIleHisPhePheGluHisAla 276
Qy 975 TTGGCATTCATCGTCTCTTACAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034
Db 277 ThrGlyIleGluAlaValLeuTyrSerProArgIlePheLysLysAlaGlyValVal 296
Qy 1035 AACGACAAACATCTTCGGCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1094
Db 297 SerLysAspLysLeuLeuLeuAlaThrValGlyValGlyLeuThrLysAlaPhePheIle 316
Qy 1095 TTGTTGGCGACTTCTTCATCGACGGCGTGGCGGGCGCGCTGCTGCTGCTGCTGCTGCTGCT 1154
Db 317 IleIleAlaThrPheLeuLeuAspLysValGlyArgArgLysLeuLeuLeuThrSerThr 336
Qy 1155 GCGGGGATATCTCTCTCATCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1214
Db 337 GlyMetValPheAlaLeuThrSerLeuAlaValSerLeuThrMetValGlnArgPhe 356
Qy 1215 CCCGACGCAACATCTGGCCCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1274
Db 357 -----GlyArgLeuAlaTrpAlaLeuSerLeuSerIleValSerThrTyrAlaPheVal 374
Qy 1275 GCTTCTTCTCATCGCTTGGCCCGCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1334
Db 375 AlaPhePheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGluIlePhePro 394
Qy 1335 CTCAGGTGCGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1394
Db 395 LeuArgLeuArgAlaGlnGlyAlaSerIleGlyValAlaValAsnArgIleMetAsnAla 414
Qy 1395 GTCACTCATGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1454
Db 415 ThrValSerMetSerPheLeuSerMetThrLysAlaIleThrThrGlyGlyValPhePhe 434
Qy 1455 CTCCTACTCGGATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1514
Db 435 ValPheAlaGlyIleAlaValAlaAlaIleTrpTrpPhePhePheMetLeuProGluThr 454

Qy 1515 CCGCGCCGACGCTGGAGAGATGACCAAGCTTTTCGCGGACACGCGCGCGCTCGGAA 1574
Db 455 LysGlyLeuProLeuGluGluMetGluLysLeuPheGlyGlyGlyProArgGlyAsp 474
Qy 1575 TCAGAC 1580
Db 475 ArgAsp 476
RESULT 13
AAG32071
ID AAG32071 standard; protein; 508 AA.
AC AAG32071;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38622.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128645P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.	Length:	508
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.	Matches:	271
PR 17-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.	Conservative:	78
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0148684P.	Mismatches:	120
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.	Indels:	13
PR 18-JUN-1999;	99US-0139456P.	PR 20-AUG-1999;	99US-0149222P.	Gaps:	3
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149723P.		
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.		
PR 18-JUN-1999;	99US-0139459P.	PR 23-AUG-1999;	99US-0149902P.		
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149930P.		
PR 18-JUN-1999;	99US-0139461P.	PR 25-AUG-1999;	99US-0150566P.		
PR 18-JUN-1999;	99US-0139462P.	PR 26-AUG-1999;	99US-0150884P.		
PR 18-JUN-1999;	99US-0139463P.	PR 27-AUG-1999;	99US-0151065P.		
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151066P.		
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.		
PR 22-JUN-1999;	99US-0139859P.	PR 30-AUG-1999;	99US-0151303P.		
PR 23-JUN-1999;	99US-0140053P.	PR 31-AUG-1999;	99US-0151438P.		
PR 23-JUN-1999;	99US-0140054P.	PR 01-SEP-1999;	99US-0151930P.		
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.		
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.		
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.		
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.		
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154038P.		
PR 01-JUL-1999;	99US-0142354P.	PR 20-SEP-1999;	99US-0154779P.		
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.		
PR 06-JUL-1999;	99US-0142290P.	PR 23-SEP-1999;	99US-0155486P.		
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.		
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.		
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156595P.		
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.		
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157533P.		
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.		
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.		
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.		
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.		
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.		
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.		
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.		
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.		
PR 20-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.		
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.		
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.		
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.		
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.		
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.		
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.		
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160768P.		
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160770P.		
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160814P.		
PR 23-JUL-1999;	99US-0145145P.	PR 21-OCT-1999;	99US-0160815P.		
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160980P.		
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160981P.		
PR 26-JUL-1999;	99US-0145276P.	PR 22-OCT-1999;	99US-0160989P.		
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161404P.		
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161405P.		
PR 27-JUL-1999;	99US-0145919P.	PR 25-OCT-1999;	99US-0161406P.		
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161359P.		
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161360P.		
PR 02-AUG-1999;	99US-0146388P.	PR 26-OCT-1999;	99US-0161361P.		
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161920P.		
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161922P.		
PR 04-AUG-1999;	99US-0147204P.	PR 28-OCT-1999;	99US-0161931P.		
PR 04-AUG-1999;	99US-0147302P.	PR 29-OCT-1999;	99US-0162142P.		
PR 05-AUG-1999;	99US-0147192P.				
PR 05-AUG-1999;	99US-0147260P.				
PR 06-AUG-1999;	99US-0147303P.				
PR 06-AUG-1999;	99US-0147416P.				
PR 09-AUG-1999;	99US-0147493P.				
PR 09-AUG-1999;	99US-0147935P.				
PR 10-AUG-1999;	99US-0148171P.				

Alignment Scores:

Pred. No.: 4.4e-96
Score: 1385.50
Percent Similarity: 72.41%
Best Local Similarity: 56.22%
Query Match: 38.34%
DB: 3

Length: 508
Matches: 271
Conservative: 78
Mismatches: 120
Indels: 13
Gaps: 3

PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134222P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-01351124P.
PR 20-MAY-1999; 99US-01353153P.
PR 21-MAY-1999; 99US-0135333P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139432P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139859P.
PR 23-JUN-1999; 99US-0140333P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145278P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 27-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.

```

PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 6,91e-90 Length: 466
Score: 1303.50 Matches: 256
Percent Similarity: 71.90% Conservatives: 74
Best Local Similarity: 55.77% Mismatches: 116
Query Match: 36.07% Indels: 13
DB: 3 Gaps: 3

US-10-051-902A-21 (1-2017) x AAG32073 (1-466)
QY 219 ATGAGCGGGCGTGTATCAAGAGGACTCAACATCAGTGACGGGAGGTGGAG 278
DB 1 MetSerGlyAlaGlnIlePheIleArgAspLeuLysIleAsnAspThrGlnIleGlu 20
QY 279 GTTCTCAGGGCATACTGAACCTTACTCGCTCATCGGCTCTTCGCGGGGGCGGACG 338
DB 21 ValLeuAlaGlyIleLeuAsnLeuCysAlaLeuValGlySerLeuThrAlaGlyLysThr 40
QY 339 TCGGACTGGATCGCGCGGTGACACATCGTGTTCGCGCGCGCTCATATCTTCGCGGG 398
DB 41 SerAspValIleGlyArgTyrThrIleAlaLeuSerAlaValIlePheLeuValGly 60
QY 399 GSGTTCCTCATGGGTTTCGCGGTCAACTAGCCCATGCTCATGTTTCGCGCGGTTCGTGGCC 458
DB 61 SerValLeuMetGlyTyrGlyProAsnTyrProValLeuMetValGlyArgCysIleAla 80
QY 459 GGCATCGCGTGGGTACGCGCTCATGATCGCGCGGTGTACACCGCGGAGGTGTCCCG 518
DB 81 GlyValGlyValGlyPheAlaLeuMetIleAlaProValTyrSerAlaGluIleSerSer 100
QY 519 CGTCGCGCGGTGCTTCGCTGACGCTGCTTCGCGGAGGTGTCACTCACTTCGCGCATCTG 578
DB 101 AlaSerHisArgGlyPheLeuThrSerLeuProGluLeuCysIleSerLeuGlyIleLeu 120
QY 579 CTCGGGTACGTCTCGAATATGCTTTCTTCGCGCTTCGCGGTGAACCTTCGCGGTGGCGCATC 638
DB 121 LeuGlyTyrValSerAsnTyrCysPheGlyLysLeuThrLeuLysLeuGlyTyrArgLeu 140
QY 639 ATGCTCGCATCGCGCGCGCGCTCGCTGCTGCTCGCGCTCATGCTGCTCGCGCATCGCG 698
DB 141 MetLeuGlyIleAlaAlaPheProSerLeuIleLeuAlaPheGlyIleThrArgMetPro 160
QY 699 GAGTCGCGCGGTGCTGCTCATGAAGGACGCGCTTCGCGGACGCAAGGCTGCTCGGAG 758
DB 161 GluSerProArgTyrPheValMetGlnGlyArgLeuGluAlaLysIleMetVal 180
QY 759 AAGACTCCGACACGCGCGGAGCGCGCGGAGCGCTGCGCGCATCAAGCGCGCGCGCC 818
DB 181 LeuValSerAsnThrGluGluAlaGluAlaGluArgPheArgAspIleLeuThrAlaAla 200
QY 819 GGCATCCCTGAGGAGCTCGACGGCGACCTGTGTGACCGTCCCAAGAGGAGGCGGAAC 878
DB 201 GluValIle-----AspValThrGluIleLysGluValGlyGlyVal 214
QY 879 GAGAGACGG-----GTGTGAAGAGGCTCATCTGTCCTCCCGACCGCGCGCC 923
DB 215 LysLysLysAsnHisGlyLysSerValTyrArgGluLeuValIleLysProArgProAla 234
QY 924 ATCGCGCGCATCTGCTGCTCGGGATCGGCATCCACATCTTCCTCCAGCATGCTTGGGCATT 983

```

```

DB 235 ValArgLeuIleLeuAlaValGlyIleHisPheGluHisAlaThrGlyIle 254
QY 984 CACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAGAGCCCGCGGATTAAACGAACGACAAA 1043
DB 255 GluAlaValValLeuTyrSerProArgIlePheLysLysAlaGlyValValSerLysAsp 274
QY 1044 CACTTCTTGGGACCACTTGGCCGTCGCTGTCACCAAGAGGCTTTTCATGTTGGCG 1103
DB 275 LysLeuLeuAlaThrValGlyValGlyLeuThrLysAlaPhePheIleIleAla 294
QY 1104 ACTTCTTCATCGAGCGGTGCGGGCGCGCGCTGTTGCTGGGAGACACGGCGGGGATA 1163
DB 295 ThrPheLeuAspLysValGlyArgArgLysLeuLeuLeuThrSerThrGlyMet 314
QY 1164 ATCTCTCTCCTCATCGGCTCGGCGCGGTACCGTCTCGGCGGACACCCGACGCC 1223
DB 315 ValPheAlaLeuThrSerLeuAlaValSerLeuThrMetValGlnArgPhe-----Gly 332
QY 1224 AGATACCTTGGCCATCGGCTTACGCTCGCTCCACCTCGCTACGTCGCTTCTTC 1283
DB 333 ArgLeuAlaTyrPheAlaSerLeuSerIleValSerThrTyrAlaPheValAlaPhePhe 352
QY 1284 TCATCGGCGCTTGGCGCGCATCACGTGGGTGTACAGCTCGAGATCTTCCGCTCCAGGTG 1343
DB 353 SerIleGlyLeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeu 372
QY 1344 CGCGGCTGGGTGCTCGCTCGGCTCGGCGCGGCGGCGGTCACCGGCGGTCTATCTCC 1403
DB 373 ArgAlaGlnGlyAlaSerIleGlyValAlaValAsnArgIleMetAsnAlaThrValSer 392
QY 1404 ATGACCTTCTCTGCTGCTGTCGCAAGGCCATCACCATCGGCGGCGGCTTCTTCTCTACTCC 1463
DB 393 MetSerPheLeuSerMetThrLysAlaIleThrThrGlyGlyValPhePheValPheAla 412
QY 1464 GCATCGCGCGCTCGCTGGGTGTCTTACACTTACCTCCCGGAGACCGCGGCGCG 1523
DB 413 GlyIleAlaValAlaAlaTyrPhePhePheMetLeuProGluThrLysGlyLeu 432
QY 1524 ACCTCGGAGGACATGCAAGCTGTTCGCGGACACGCGCGCGCTCGGAATCAGAC 1580
DB 433 ProLeuGluGluMetGluLysLeuPheGlyGlyGlyProArgGlyAspArgAsp 451

RESULT 15
ADA48320
ID ADA48320 standard; protein; 333 AA.
XX
AC ADA48320;
XX
XX 20-NOV-2003 (first entry)
DE Rice protein conferring disease resistance in plants.
XX disease resistance; pathogen tolerance; plant pathogen; plant; rice.
XX Oryza sativa.
OS
XX WO2003000906-A2.
XX
XX 03-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-IB002453.
PF
XX
XX 22-JUN-2001; 2001US-0300112P.
PR
XX 25-SEP-2001; 2001US-032277P.
PR
XX 22-MAR-2002; 2002US-0366535P.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA
XX Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
PI Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
XX WPI; 2003-184052/18.
DR

```


This Page Blank (uspto)

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:42:55 ; Search time 24 Seconds
(without alignments)
8677.465 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 3614
Sequence: 1 cttacatgtaagctcgtgc.....aaaaaaaaaaaaaaaaaaaaa 2017

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 77828

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USTO_spool_p/US10051902/runat 30062004_164724_20438/app query.fasta_1.2183
-DB=Issued Patents AA -QMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051902 @CGN 1 1 27 @runat 30062004_164724_20438 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIME=30 -HEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	2570	71.1	510	US-09-291-922-22
2	2015	55.8	539	US-09-291-922-22
3	1877.5	52.0	529	US-09-291-922-28
4	1872.5	51.8	513	US-09-291-922-20
5	1660	45.9	523	US-09-291-922-24
6	1589	44.0	549	US-09-291-922-30
7	682	18.9	167	US-09-291-922-18
8	672	18.6	488	US-10-162-012-46
9	591	16.4	514	US-09-489-039A-11902
10	585	16.2	476	US-09-489-039A-11933
11	576	15.9	584	US-08-928-692-13
12	576	15.9	584	US-09-339-972-13

13	549	15.2	517	4	US-09-679-686B-18	Sequence 18, Appl
14	536	14.8	514	4	US-09-679-686B-22	Sequence 22, Appl
15	533	14.7	502	4	US-09-679-686B-2	Sequence 2, Appl
16	520	14.4	511	4	US-09-679-686B-12	Sequence 12, Appl
17	514.5	14.2	501	4	US-09-489-039A-11731	Sequence 11731, A
18	507.5	14.0	519	4	US-09-679-686B-24	Sequence 24, Appl
19	505.5	14.0	510	4	US-09-679-686B-19	Sequence 19, Appl
20	500	13.8	729	4	US-09-291-922-29	Sequence 29, Appl
21	499	13.8	518	4	US-09-679-686B-23	Sequence 23, Appl
22	489.5	13.5	504	4	US-09-679-686B-21	Sequence 21, Appl
23	487.5	13.5	518	4	US-09-679-686B-16	Sequence 16, Appl
24	487	13.5	562	4	US-10-162-012-44	Sequence 44, Appl
25	481	13.3	737	4	US-09-291-922-8	Sequence 8, Appl
26	469	13.0	493	3	US-09-031-392-10	Sequence 10, Appl
27	469	13.0	493	3	US-09-299-549-10	Sequence 10, Appl
28	469	13.0	493	4	US-09-610-417-10	Sequence 10, Appl
29	453	12.5	521	4	US-09-489-039A-9549	Sequence 9549, Ap
30	448	12.4	509	2	US-09-031-392-6	Sequence 6, Appl
31	448	12.4	509	3	US-09-299-549-6	Sequence 6, Appl
32	448	12.4	509	4	US-09-610-417-6	Sequence 6, Appl
33	443	12.3	494	2	US-09-031-392-5	Sequence 5, Appl
34	443	12.3	494	3	US-09-299-549-5	Sequence 5, Appl
35	443	12.3	494	4	US-09-610-417-5	Sequence 5, Appl
36	439.5	12.2	492	2	US-08-358-844-3	Sequence 3, Appl
37	439.5	12.2	492	5	PCT-US95-16126-3	Sequence 3, Appl
38	438.5	12.1	488	2	US-08-928-692-11	Sequence 11, Appl
39	438.5	12.1	488	4	US-09-339-972-11	Sequence 11, Appl
40	437	12.1	747	4	US-09-291-922-2	Sequence 2, Appl
41	434.5	12.0	323	4	US-09-134-001C-4635	Sequence 4635, Ap
42	422	11.7	524	2	US-08-928-692-12	Sequence 12, Appl
43	422	11.7	524	4	US-09-339-972-12	Sequence 12, Appl
44	420.5	11.6	534	2	US-09-031-392-4	Sequence 4, Appl
45	420.5	11.6	534	3	US-09-299-549-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-291-922-22 ; Sequence 22, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
US-09-291-922-22

Alignment Scores:
Pred. No.: 3,19e-189 Length: 510
Score: 2570.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 71.11% Indels: 0
DB: 4 Gaps: 0

US-10-051-902A-21 (1-2017) x US-09-291-922-22 (1-510)

Db 46 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyThrAspIleGly 65
QY 216 GTGATGAGCGGGCGTGGTGTACATCAAGAGGACTTCAACATCAGTACCGGAGGTG 275
Db 66 ValMetSerGlyAlaSerLeuTyrlleGlnLysAspLeuLysIleAsnAspThrGlnLeu 85
QY 276 GAGTTTCTCATGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGGGGGG 335
Db 86 GluValLeuMetGlyIleLeuAsnValTySerLeuIleGlySerPheAlaIleGlyArg 105
QY 336 ACCTCGGACTGATCGCGCGGGTACACATCGTGTTCGCGCGCTCATATCTTCGCG 395
Db 106 ThrSerAspTrpIleGlyArgPheThrIleValPheAlaValIlePhePheAla 125
QY 396 GGGGGTTCTCTCATGGGTTCGGCGTCACTACGCGTCTCATGTCGCGCTTCGCG 455
Db 126 GlyAlaLeuIleMetGlyPheSerValAsnTyAlaMetLeuMetPheGlyArgPheVal 145
QY 456 GCCGGCATCGCGTGGGCTACGGGCTCATGATCGCGCGGTGTACACCGCGAGGTGTCG 515
Db 146 AlaGlyIleGlyValGlyTyAlaLeuMetIleAlaProValAsnThrGlyGluValSer 165
QY 516 CCGCGTTCGGCGGTGCTTCAGTGTTCGCGAGGTGTTCATCAACTTCGCGATC 575
Db 166 ProAlaSerAlaArgGlyValLeuThrSerPheProGluValPheIleAsnPheGlyIle 185
QY 576 CTGCTCGGTGACGTCTCGAACTATGTTCTTCGCGTTCGCGTGAACCTCGGTCGCGC 635
Db 186 LeuLeuGlyTyValSerAsnPheAlaPheAlaArgLeuSerLeuArgIleGlyTrpArg 205
QY 636 ATCATGCTCGGATCGCGCGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 695
Db 206 IleMetLeuGlyIleGlyAlaValProSerValLeuLeuAlaPheMetValLeuGlyMet 225
QY 696 CCGAGTTCGCGCGGTGCTGTCATGAAGGAGCGCTCGCGAGCGCGTTCGCGAGGTGTCG 755
Db 226 ProGlnSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 245
QY 756 GAGAAGACCTCCGACAGCGGAGGCGCGCGGCGCTCGCGACCATCAAGCGCGCC 815
Db 246 AlaLysThrSerAspThrProGluGluAlaGluArgIleAlaAspIleLysThrAla 265
QY 816 CCGCGCATCTCGAGGAGCTCGAGCGGAGTGTGACCGTCCCGAG---AGAGGGAGC 872
Db 266 AlaGlyIleProLeuGlyLeuAspGlyAspValProValProLysAsnLysGlySer 285
QY 873 GGAACGAGAGCGGTGTGGAGGAGCTCATCTGTCCCGACCGCGCGCATCGCGCGC 932
Db 286 SerGluGlnLysArgValLeuLysAspLeuLeuSerProThrIleAlaMetArgHis 305
QY 933 ATCTGTGTTCGGGATCGGATCCACTTCTTCAGCATCGGTGGCATCTCATCTCGTC 992
Db 306 IleLeuIleAlaGlyIleGlyIleHisPhePheGlnGlnSerSerGlyIleAspAlaVal 325
QY 993 GTCTTCTACGCCCTCTCGTGTTCAGAGCCCGGATTACGACGACCAACACTTCTTG 1052
Db 326 ValLeuTySerProLeuValPheLysSerAlaGlyIleThrGlyAspSerArgLeuArg 345
QY 1053 GGCACCACTTGGCGGTTCGTGTCCACAGAGGCTTTTCATCTGTGTGGCATCTTCTTC 1112
Db 346 GlyThrThrValAlaValGlyAlaThrAsnThrValPheIleLeuValAlaThrPheLeu 365
QY 1113 ATCAGCGCTCGCGCGCGCGCTGTGTGGAGACGACGCGCGGATAATCTCTCC 1172
Db 366 LeuAspArgIleArgArgProLeuValLeuThrSerThrGlyGlyMetLeuValSer 385
QY 1173 CTCACTCGGCTCGCGCGCGGCTCACCGTGTTCGCGGACGACCGCGCGCAAGATACCT 1232
Db 386 LeuValGlyLeuAlaThrGlyLeuThrValIleSerArgHisProAspGluLysIleThr 405
QY 1233 TGGGCGATCGGCTTAAGCATCGCTCCACCGCTCGCTACGTCGCTTCTCTCCATCGC 1292

Db 406 TrpAlaIleValLeuCysIlePheCysIleMetAlaTyValAlaPhePheSerIleGly 425
QY 1293 CTTGGCCCATCAGTGGTGTACAGTCCGAGATCTTCCCGCTCCAGGTCCGCGCGGTG 1352
Db 426 LeuGlyProIleThrTrpValTySerSerGluIlePheProLeuHisValArgAlaLeu 445
QY 1353 GCGTCTCGTTCGCGCGTCCGCGCAACCGCGTCAACGCGCGGTCACTCTCCATGACCTTC 1412
Db 446 GlyCysSerLeuGlyValAlaValAsnArgLeuThrSerGlyValIleSerMetThrPhe 465
QY 1413 CTGTCGCTCCAAAGGCATCACCATCGCGCGAGTCTTCTCTCTACTCCGGCATCGCC 1472
Db 466 IleSerLeuSerLysAlaMetThrIleGlyAlaPhePheLeuPheAlaGlyIleAla 485
QY 1473 GCGTCTCGCTGGGTGTTCTTCTACACCTTACCTCCCGGAGACCCCGCGGACGTGGAG 1532
Db 486 SerPheAlaTrpValPhePheAlaTyLeuProGluThrArgGlyArgThrLeuGlu 505
QY 1533 GAGATGACCAAGCTGTTCCGGCGACACGGCC-----GCCGCTCGGAATCAGAC 1580
Db 506 AspMetSerSerLeuPheGlyAsnThrAlaThrHisLysGlnGlyAlaAlaGluAlaAsp 525
QY 1581 GAGCCAGCCAAAGGAGAGAGAGGTGGAATGGCCGCACTAAC 1625
Db 526 AspAspAlaGlyGlu---LysLysValGluMetAlaAlaThrAsn 539
RESULT 3
US-09-291-922-28
; Sequence 28, Application US/0929:922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Ritz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-28
Alignment Scores:
Pred. No.: 4,8e-136 Length: 529
Score: 1977,50 Matches: 369
Percent Similarity: 81,38% Conservative: 55
Best Local Similarity: 70,83% Mismatches: 92
Query Match: 51,95% Indels: 5
DB: 4 Gaps: 3
US-10-051-902A-21 (1-2017) x US-09-291-922-28 (1-529)
QY 60 CCAGAGCGA-----GCCCTCTCTCTCTGACACCGGAGATGGCTTCGCGCGGTG 113
Db 5 ProGluArgLysGlyAlaGluAspLysGluGlySerArgMetAlaSerAlaAlaLeu 24
QY 114 CCGGAG-----GCCGTCGCGCGGAGAGAGGCAACGTCGCTTCGCTTCGCTTC 167
Db 25 ProGluProGlyAlaValHisProArgAsnLysGlyAsnPheLysTyAlaPheThrCys 44
QY 168 GGCATCTCGCTCCATGACCTTCCTCTCGGCTACGATATCGGGGTGATGACGGG 227
Db 45 AlaLeuCysAlaSerMetAlaThrIleValLeuGlyTyAspValGlyValMetSerGly 64
QY 228 GCGTCGCTGTACATCAAGAGGACTTCAACATCAGTACGCGGAGGTGAGGTTCTCATG 287

Db 65 AlaSerLeuTyrIleLysArgAspLeuGlnIleThrAspValGlnLeuGluIleMetMet 84
QY 288 GGCATACCTCTACTCGCTCATCGCTCTTCGGCGGGCGGAGCGATCGGAGTGG 347
Db 85 GlyIleLeuSerValTyrAlaLeuIleGlySerPheLeuGlyAlaArgThrSerAspTrp 104
QY 348 ATCGCGCGGCGTACACCATGTTGTCGCCCGCTCATATTCTTCGGCGGGGSGTTCCTC 407
Db 105 ValGlyArgArgValThrValValPheAlaAlaIlePheAsnAsnGlySerLeuLeu 124
QY 408 ATGGGTTTCGGCGTCAACTACGCTCATGTTTCGGCGCTTCGTCGGCGGATCGGC 467
Db 125 MetGlyPheAlaValAsnTyrAlaMetLeuMetValGlyArgPheValThrGlyIleGly 144
QY 468 GTGGGCTACGGCTCATGATCGCGCGGTGTACACCGCGGAGGTGTTCGGCGGCTCGCG 527
Db 145 ValGlyTyrAlaIleMetValAlaProValTyrThrProGluValSerProAlaSerAla 164
QY 528 CGTGGCTTCGAGCTGCTTCGGAGGTGTTCATCACTTCGCGATCCTGCTCGGCTAC 587
Db 165 ArgGlyPheLeuThrSerPheThrGluValPheIleAsnValGlyIleLeuLeuGlyTyr 184
QY 588 GTCTGGAATACTGCTTCCTCCGCTTCGCGTGAACCTCGGTCGCGCATCATGCTCGGC 647
Db 185 ValSerAsnTyrAlaPheAlaArgLeuProLeuHisLeuSerTyrArgValMetLeuGly 204
QY 648 ATCGCGCGGCGCGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707
Db 205 IleGlyAlaValProSerAlaLeuLeuLeuAlaMetValPheGlyMetProGluSerPro 224
QY 708 CGTGGCTGCTGTCATGAGGAGCGCTCGCGGACCCCAAGGTGTGTGGAAGACATCC 767
Db 225 ArgTrpLeuValMetLysGlyArgLeuAlaAspAlaArgAlaValLeuAlaLysThrSer 244
QY 768 GACAGCGGAGGAGCGCGGAGCGCTGCGGACATCAAGCGCGCGCGCGGATCCCT 827
Db 245 AspThrProGluGluAlaValGluArgLeuAspGlnIleLysAlaAlaGlyIlePro 264
QY 828 GAGGAGCTCGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 887
Db 265 ArgGluLeuAspGlyAspValValMetProLysThrLysGlyGlnGluLysGln 284
QY 888 GTGTGGAAGAGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 947
Db 285 ValTrpLysGluLeuIlePheSerProThrProAlaMetArgArgIleLeuLeuAlaAla 304
QY 948 ATCGGATCCATCTTCCAGCATGCTTGGGATTCACCTCGCTGCTTCTACAGCCTT 1007
Db 305 LeuGlyIleHisPhePheGlnAlaThrGlySerAspSerValValLeuTyrSerPro 324
QY 1008 CTCGTGTTCAAGAGCCCGGATTAAACGACGACAAACACTTCTTGGGACCATTTGGCCG 1067
Db 325 ArgValPheGlnSerAlaGlyIleThrGlyAspAsnHisLeuLeuGlyAlaThrCysAla 344
QY 1068 TTCGTGTCCACAGAGCTTTTCATCTTGTGGGATTTCTTCATCGAGCGCTGGG 1127
Db 345 MetGlyValMetLysThrLeuPheIleLeuValAlaThrPheGlnLeuAspArgValGly 364
QY 1128 CGGCGGCGCTGTTGCTGGGCGAGCAGCGGCGGATATCTCTTCCTCATCGGCTCGGC 1187
Db 365 ArgArgProLeuLeuLeuThrSerThrAlaGlyMetLeuAlaCysLeuIleGlyLeuGly 384
QY 1188 GCGGGCTACCGTGTGCGGCGAGACCCCGACGCGCAAGATACCTTGGGCGATCGGCTA 1247
Db 385 ThrGlyLeuThrValValGlyArgHisProAspAlaLysValProTrpAlaIleGlyLeu 404
QY 1248 AGATCCCTCCACCTTCGCTAGTGGCTTCTTCTCCATCGGCTTGGCGCCATCAG 1307
Db 405 CysIleValSerIleLeuAlaTyrValSerPhePheSerIleGlyLeuGlyProLeuThr 424
QY 1308 TGGGTGTACAGTCTGGAGATCTTCCCGCTCCAGTGGCGGCTGGGCTGCTGCTGCTG 1367
Db 425 SerValTyrThrSerGluValPheProLeuArgValAlaAlaLeuGlyPheAlaLeuGly 444

QY 1368 GTCCGCGCAACCGGTCACCGCGCTCATCTCATGACCTTCTGCTGCTGCTGCTGCTG 1427
Db 445 ThrSerCysAsnArgValThrSerAlaAlaValSerMetSerPheLeuSerLeuSerLys 464
QY 1428 GCATCACCATCGCGGCGAGCTTCTCTCTACTCGGATCGCGCGCTCGCTCGGCTG 1487
Db 465 AlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyIleAlaAlaIleGlyTrpIle 484
QY 1488 TTCTTCTACACCTACTCTCCGCGAGACCCGCGCGGACGCTCGAGGATGAGCAAGTGT 1547
Db 485 PhePhePheThrPheIleProGluThrArgGlyLeuProLeuGluGluIleGlyLysLeu 504
QY 1548 TTCGCGGACACCGCGCGCTCGGAAATCAGACGAGCCAGCC---AAGAGAGAGAGAG 1604
Db 505 PheGlyMetThrAspThrAlaValGluAlaGlnAspThrAlaThrLysAspLysAlaLys 524
QY 1605 GTG 1607
Db 525 Val 525
RESULT 4
US-09-291-922-20
; Sequence 20, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-09-291-922-20
Alignment Scores:
Pred. No.: 1,15e-135 Length: 513
Score: 1872.50 Matches: 372
Percent Similarity: 84.38% Conservative: 60
Best Local Similarity: 72.66% Mismatches: 77
Query Match: 51.81% Indels: 3
DB: 4 Gaps: 2
US-10-051-902A-21 (1-2017) x US-09-291-922-20 (1-513)

QY 96 ATGGCTTCGCGCGCTCGCGAGCGCTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 155
Db 1 MetAlaSerAspGluLeuAlaLysAlaValGluProArgLysLysGlyAsnValLysTyr 20
QY 156 GCCTTCGCTCGCGCATCTCGCTCCATGACCTCCATCCCTCGCTCGCTCGCTCGCTCG 215
Db 21 AlaSerIleCysAlaIleLeuAlaSerMetAlaSerValIleLeuGlyTyrAspIleGly 40
QY 216 GTGATGAGCGGGCGCTGCTGTATCATCAAGAGACTTCAACATCATGAGCGGAGAGGTG 275
Db 41 ValMetSerGlyAlaAlaMetTyrIleLysLysAspLeuAsnIleThrAspValGlnLeu 60
QY 276 GAGGTTCTCATGGGCATCATCACTCTACTCGCTCATCGCTCGCTCGCTCGCTCGCT 335
Db 61 GluIleLeuIleGlyIleLeuSerLeuTyrSerLeuPheGlySerPheAlaGlyAlaArg 80
QY 336 ACCTCGGACTGGATCGCGCGCGCTGATCACCATCGTGTTCGCGCGCTGCTGCTGCTG 395
Db 81 ThrSerAspArgIleGlyArgArgLeuThrValValPheAlaAlaValIlePhePheVal 100


```
QY 519 GCGTCGGCGGCTGCTTCCTGACGTCGTTCCCGGAGGTGTTTCATCAACTCCGCGCATCTG 578
Db 152 AlaSerArgGlyPheLeuThrSerPheProGluValPheIleAsnGlyGlyIleLeu 171
QY 579 CTCGGGTACGCTCGAACTATGCTTCTCCGCTTCCGCTGGAACCTCCGGTGGCGCATC 638
Db 172 IleGlyTyrIleSerAsnTyrAlaPheSerIysLeuThrLeuIysValGlyTyrArgMet 191
QY 639 ATGCTCGGCATCGCGGCGCCCTCCGCTGCTGCTCGCGCTCATGTCGTGCTCGCATCCG 698
Db 192 MetLeuGlyValGlyAlaIleProSerValLeuThrValGlyValLeuAlaMetPro 211
QY 699 GAGTCGCGCGGTGCTGCTCATCAAGGAGCGCTCCGCGACGCCAAGTGGTCTGGAG 758
Db 212 GluSerProArgTyrPheValMetArgGlyArgLeuGlyGluAlaArgLysValLeuAsn 231
QY 759 AAGACCTCCGACACGCGGAGGAGCGCGGAGCGCTGCGCGACATCAAGCGCGCGCC 818
Db 232 LysThrSerAspSerLysGluGluAlaGlnLeuArgLeuAlaGluIleLysGlnAlaAla 251
QY 819 GGCATCCTGAGGAGCTCGACGCGACGCTGTCACCTCCCGACAGAGGAGCGGAAAC 878
Db 252 GlyIleProGluSerCysAsnAspValValGlnValAsnLysGlnSerAsnGly--- 270
QY 879 GAGAAGCGGTGTGGAAGAGGCTCATCTCTCCCGACCCCGCGCATCGCGCGCATCTG 938
Db 271 --GluGlyValTyrLysGluLeuPheLeuTyrProThrProAlaIleArgHisIleVal 289
QY 939 CTGTCGGGATCGCATCCACTTCTCCAGACGCTGTCGCTGGCGATTCCTCCGCTGCTTC 998
Db 290 IleAlaAlaLeuGlyIleHisPhePheGlnGlnAlaSerGlyValAspAlaValLeu 309
QY 999 TACAGCCTCTCGTGTCAAGAGCCCGGATTACAGAACGACAACTCTCTTGGGCACC 1058
Db 310 TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuAla 329
QY 1059 ACTTGGCGGTGCTGTCACAAAGAGCTTTTCATCTTGTGGCGACTTCTTTCATCGAC 1118
Db 330 ThrValAlaValGlyPheValLysThrValPheIleLeuAlaAlaThrPheThrLeuAsp 349
QY 1119 GGCCTCGCGGCGCGCTGCTGCTGGGACGACGCGCGGATATCTCTCCCTCATC 1178
Db 350 ArgValGlyArgProLeuLeuLeuSerValGlyGlyMetValLeuSerLeuLeu 369
QY 1179 GGCCTCGCGGCGGCTCACCGTCTCGGCGACGACCCCGACCAAGATACCTCGGGCC 1238
Db 370 ThrLeuAlaIleSerLeuThrValIle--AspHisSerGluArgLysLeuMetTyrAla 388
QY -1239 ATCGGCTAAGCATCGCTCCACCTCGCTAGTCGCTCTTCTCCATCGCGCTGGC 1298
Db 389 ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly 408
QY 1299 CCCATCACGTGGGTGACAGCTCGGAGATCTTCCGCTCCAGGTGGCGGCTGGCTGC 1358
Db 409 ProIleThrTyrValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnGlyAla 428
QY 1359 TCGCTCGGCTGCGCGCCACCGGTCACCGGCGGTATCTCCATGACCTTCTCTGTCG 1418
Db 429 AlaAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448
QY 1419 CTGTCGAAGCCATCACATCGCGGCGAGCTTCTCTCTACTCCGCGCATCGCGCGCTC 1478
Db 449 LeuThrArgAlaIleThrIleGlyGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal 468
QY 1479 GCCTGGGTGTTCTTACACTACCTACCTCCGAGACCCGCGCGGAGCGCTGGAGGAGATG 1538
Db 469 GlyTyrPhePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488
QY 1539 AGCAAGCTGTTCGGCGACACGCGCGCGCTCGGAATCAGACGAGCCGACCAAGGAGAG 1598
Db 489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla-----Ser 502
QY 1599 AAGAAGGTGGAATGCGCGCCACTAACTGATCAACTAACCGCAAAATCACCAAATCTCTTA 1658
```

```
Db 503 LysAlaValGlu-----AsnGluAsnGlyGlnValAla 513
QY 1659 AGGGTTCCTTCGCAAAACAGTGTG 1682
Db 514 GlnValGlnLeuGlyThrAsnVal 521

RESULT 6
US-09-291-922-30
; Sequence 30, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-291-922-30

Alignment Scores:
Pred. No.: 6,89e-114 Length: 549
Score: 1589.00 Matches: 326
Percent Similarity: 72.23% Conservative: 59
Best Local Similarity: 61.16% Mismatches: 126
Query Match: 43.97% Indels: 22
DB: 4 Gaps: 5

US-10-051-902A-21 (1-2017) x US-09-291-922-30 (1-549)
```

```
QY 78 TCCCTGCACACACGAGATGGCTTCGCCGCGCTG-----CCGAGGCGGTC 125
Db 10 SerAspProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 29
QY 126 GCGCGAAGAAAGGCAACGCTCCGCTTCGCCCTTCGCCCATCCTCGCTCCATG 185
Db 30 ProProlLysArgAsn-----LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 46
QY 186 ACCTCCATCTCTCGGCTAGCATATCGGGGTGATAGCGGGGCGTGCCTGTACATCAAG 245
Db 47 ThrSerValLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaIleIleTyrLeuLys 66
QY 246 AAGACTTCAACATCAGTCAGCGAAGTGAGGTTCATGGGCATACTGAACCTCTAC 305
Db 67 GluAspThrHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 86
QY 306 TCGTCTATCGGCTCTCTCGCGCGGCGGACGTCGAGCTGATCGCGCGGCTACACC 365
Db 87 CysLeuPheGlySerPheAlaAlaGlyArgThrSerAspTrpIleGlyArgTyrThr 106
QY 366 ATCGGTTCGCGCGCTCATATTCCTCGCGGGSGGTTCCTCATGGGTTCGCGCTCAAC 425
Db 107 IleValLeuAlaGlyAlaIlePhePheValGlyValGlyValGlyValGlyValGly 126
QY 426 TACGCCATGCTCATGTTCCGCGCTTCGCGCGGATCGCGGTGGGCTACGCGCTCATG 485
Db 127 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 146
QY 486 ATCGCGCGGTGTACACGCGGAGGTTCGCGCGGCTCGCGGTGCTTCCTGACGTG 545
Db 147 IleAlaProValTyrThrAlaGluValSerProAlaSerArgGlyPheLeuSer 166
QY 546 TTCCCGGAGGTGTTTCATCAACTTCGCGCATCTCGCTCGGGTACGTCTCGAACTGCTTTC 605
```



```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
290 GlyIleAsnAlaIlePheTyrTyrSerProThrIlePheIleValGlyValSerAsp 309
1038 GACAAACACITCTTGGGACACACATGGCCGCTCGTGTCCACCAAGAGCGCTTTCATCTTG 1097
310 SerValAlaSerLeuLeuValThrIleIleValGlyValValAsnPheValPheThrPhe 329
1098 TTGCGC---ACTTCTTCATCAGCAGCGGTGCGGGCGCGCGCTGTCTCTGGCAGCAGC 1154
330 ValAlaLeuIlePheLeuValAspArgPheGlyArgProLeuLeuLeuGlyAla 349
1155 GGCGGGATATCTCTCCCTCATCGGCTCGCGCGCGCGCTACCGTCGTC----- 1205
350 AlaGlyMetAlaIleCysPheLeuIleLeuGlyAlaSerIleGlyValAlaLeuLeu 369
1206 -----GGCCAGCACCCCGACGCCCAAGATACCTTGGGCCATCGGCCTA---AGCATC 1253
370 LeuAsnIysProIysAspProSerSerIys-----AlaAlaGlyIleValAlaIle 386
1254 GCCTCCACCTCGGCTAGTCGCTCTCTTCCATCGCGCTTGGCCCATCAGCTGGGTG 1313
387 ValPheIleLeuLeuPheIleAlaPhePheAlaLeuGlyTrpGlyProIleProTrpVal 406
1314 TACAGCTCGGAGATCTCCGCTCCAGTGGCGCGCTGGGCTCTCGCTCGGCGCTGCC 1373
407 IleLeuSerGluLeuLeuPheProThrLysValArgSerLysAlaLeuAlaLeuAlaThrAla 426
1374 GCCAACCGCGTACCAGCGGCGTCATCTCCATGACCTTCTCTCGCTGTCCAAAGGCCATC 1433
427 AlaAsnTrpLeuAlaAsnPheIleGlyPheLeuPheProTyrIleThrGlyAlaIle 446
1434 -----ACCATCGGGGAGCTTCTTCTCTC---TACTCGGCATCGCGGCTCGCTCG 1484
447 GlyLeuAlaLeuGlyGlyTyrValPheLeuValPheAlaGlyLeuLeuValLeuPheIle 466
1485 GTGTCTCTACACTACTCTCCCGGAGACCGCGCGCGAGCGCTGGAGGAGATCAGCAAG 1544
467 LeuPheValPhePhePheValProGluThrLysGlyArgThrLeuGluGluIleGluGlu 486
1545 CTGTTC 1550
|||||
487 LeuPhe 488

RESULT 9
US-09-489-039A-11902
; Sequence 11902, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBS
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11902
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11902

```

Alignment Scores:	
Pred. No.:	2.95e-37
Score:	591.00
Percent Similarity:	51.82%
Best Local Similarity:	29.76%
Query Match:	16.35%
DB:	4
Length:	11
Matches:	14
Conservative:	109
Mismatches:	194
Indels:	44
Gaps:	11

US-10-051-902A-21 (1-2017) x US-09-489-039A-11902 (1-514)

Qy	93	GAGATGGCTTCGGCGCGCTCCGGAGGCGCTCGCGCGAAG---AGAAGGGCAACGTC	149
Db	41	LysMetThrSerIleSerAsnAspSerThrLeuSerProArgThrGlnArg	60
Qy	150	CGERTCCCTTCGGCTCGGCATCCTCGCCTCCATGACCTCCATCTCCTCGGTACCAT	209
Db	61	ArgMetAsnTrpPheValSerIleAlaAlaValAlaGlyLeuLeuPheGlyLeuAsp	80
Qy	210	ATCGGGGTGATACGGGGCTCGCTGTACATCAAGAAGGACTTCAACATCAGTCACGGG	269
Db	81	IleGlyValIleSerGlyAlaLeuProPheIleThrAspHisPheThrLeuSerSerGln	100
Qy	270	AAGGTGAGGTTTCATGGGCATACTGAACCTCTACTCGCTCATCGCTCTCTCGCGCG	329
Db	101	LeuGlnGluTrpValIleSerSerMetLeuGlyAlaAlaIleGlyAlaLeuPheAsn	120
Qy	330	GGCGGACGTTCGGAGTCGGCGCGCGGTACACCATCGGTTCGGCGCGCTCATATTC	389
Db	121	GlyTrpLeuSerPheArgLeuGlyArgIlySerLeuMetAlaGlyAlaValLeuPhe	140
Qy	390	TTCGGGGGGTTCCTCATGGGTTCCGCTCAACTACGCCATGCTCATGTTCTGGCGCG	449
Db	141	ValAlaGlySerIleGlySerAlaPheAlaAlaSerValGluValLeuLeuValAlaArg	160
Qy	450	TTCGTGGCGGCATCGCGCTAGCGCTCATGATCGCGCGCGGTGTACACCGCCGAG	509
Db	161	ValValLeuGlyValAlaValGlyLeuAlaSerTyrThrAlaProLeuTyrLeuSerGlu	180
Qy	510	GTGTCGGCGCGTTCGGCGGTCTCTGAGCTGTTCCCGGAGGTTCATCAACTTC	569
Db	181	MetAlaSerGluAsnValArgGlyLeuMetIleSerMetTyrGlnLeuMetValThrLeu	200
Qy	570	GGCATCTGTCGGGTACGCTCGAACTATGCTTTCTCCGCTTCGCCGCTGAACCTCGG	629
Db	201	GlyIleValLeuAlaPheLeuSerAspThrAlaPheSer-----TyrSerGlyAsn	217
Qy	630	TGGCGCATCATGTCGGCATCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTC	689
Db	218	TrpArgAlaMetLeuGlyValLeuAlaLeuProAlaValIleLeuIleIleLeuValVal	237
Qy	690	GGCATCGCGAGTCGCGCGTGGTGGTATGATGAGGAGCGCTCGCGGAGCCCAAGTG	749
Db	238	PheLeuProAsnSerProArgTrpLeuAlaGluIlyGlyArgHisIleGlyAlaGluGlu	257
Qy	750	GTGCTGAGGAAGACCTCCGACACGGCGGAGGAGCGCGGAGCGCTGCGCGCATCAAG	809
Db	258	ValLeuArgMetLeuArgAspThrSerGluLysAlaArgAspGluLeuAsnGluIleArg	277
Qy	810	GCOCGCCCGCATCCTCGAGGACTCGACGGCGACGTGTGACGCTCCCAAGAGAGGG	869
Db	278	GluSerLeuLysLeu-----LysGlnGly	285
Qy	870	AGCGGAACAGAGCGGTGTGAAGGAGCTCATCTGCTCCCGACCCCGCGCATCGG	929
Db	286	Gly-----Trp-----AlaLeuPheLysIleAsnArgAsnValArg	297
Qy	930	CGCATCTGTCGCGGATCGGCATCATCTTCTTCAGCATCGTGTGGGATTCACCTCC	989
Db	298	ArgAlaValPheLeuGlyMetLeuGlnAlaMetGlnPheThrGlnPheThrGlyMetAsnIle	317
Qy	990	GTGCTGCTACAGCCCTCTGCTTCAAGAGCCCGGATTAAAGCAACACACATC	1049
Db	318	IleMetTyrTrpAlaProArgIlePheLysMetAlaGlyPheThrThrThrGluGlnGln	337
Qy	1050	TGCGGACCATCTGGCGCTTCGGGTCCACCAAGAGGCTTTTCATCTTGTTCGGCACTTC	1109
Db	338	MetIleAlaThrLeuValValGlyLeuThrPheMetPheAlaThrPheIleAlaValPhe	357
Qy	1110	TTCATCAGCGGCTCGCGCGCGCGCTGTGTGTCGGCAGCACGGCGGGATATATCCTC	1169
Db	358	ThrValAspLysAlaGlyArgLysProAlaLeu-----LysIleGlyPheSerValMet	375
Qy	1170	TCCCTCATCGGCTCGCGCGCGGCTCACCGCTCGCGCCAG-----	1219

```

Db 376 AlaLeu-----GlyThrLeuValLeuGlyTyrCysLeuMetGlnPheAsp 390
QY 1212 -----CACCCGAGCCAGATACCTTGGCCATCGGCTAAGCATCGCCTCACCCCTC 1265
Db 391 AsnGlyThrAlaSerSerGlyLeuSerTrp-----LeuSerValGlyMetThrMet 407
QY 1266 GCGTACGTCGCTTCTCTCATCGGCTTGGCCCATCATCGTGGGTGTACAGTCCGAG 1325
Db 408 MetCysIleAlaGlyTyrAlaMetSerAlaAlaProValTrpIleLeuCysSerGlu 427
QY 1326 ATCTTCGCGTCCAGTGGCGGCGCTCGCTGCTCGGCTCGGCGTCCGCGCACCGGTC 1385
Db 428 IleGlnProLeuLysCysArgAspPheGlyIleThrCysSerThrThrThrAsnTrpVal 447
QY 1386 ACCAGCGGCTCATCTCATGACCTTCTCTGCTGCTGCCAGGCTCCACCATCGGCGGC 1445
Db 448 SerAsnMetIleIleGlyAlaThrPheLeuThrLeuLeuAspIleIleGlyAlaAlaGly 467
QY 1446 AGTCTTCTCTCTACTCCGGATC---GCCGCGCTCGCCTGGGTGTCTTCTACACCTAC 1502
Db 468 ThrPheTrpLeuTyrThrAlaLeuAsnValAlaPheIleIleThrPheTrp---Leu 486
QY 1503 CTCGCGGAGACCGCGCGGAGCTCGAGGATGACGAAG 1544
Db 487 IleProGluThrLysAsnValThrLeuGluHisIleGluArg 500

RESULT 10
US-09-489-039A-11933
; Sequence 11933, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11933
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11933

Alignment Scores:
Pred. No.: 7,99e-37 Length: 476
Score: 585.00 Matches: 146
Percent Similarity: 49.49% Conservative: 97
Best Local Similarity: 29.74% Mismatches: 212
Query Match: 16.19% Indels: 36
DB: 4 Gaps: 8

US-10-051-902A-21 (1-2017) x US-09-489-039A-11933 (1-476)
QY 90 CCGGAGATGGTTCGCGCGCTCCGCGAGCGCGTCCGCGCG-----AAGAAGAAGGCG 143
Db 1 ProSerThrGluSerIleThrGlnLeuGluGlyValMetProAspAsnLysGlnGly 20
QY 144 -----AACGTCGCGTTCGCTTCGCTCGCCATCTCGCTCGCTCCATGACCTCCATCCTC 197
Db 21 ArgSerAsnLysThrMetThrPhePheValCysPheLeuAlaAlaLeuAlaGlyLeuLeu 40
QY 198 CTCGGCTACGATACGGGTGATAGCGGGCGTGTGTATCATCAAGAGAGACTTCAAC 257
Db 41 PheGlyLeuAspIleGlyValIleAlaGlyAlaLeuProPheIleAlaAsnGluPheGln 60
QY 258 ATCAGTCACGGGAAGGTGGAGGTCTCTCATGGGCATCTGAACCTCTACTCGCTCATCGGC 317
Db 61 IleSerAlaHisThrGlnGluTrpValValSerSerMetMetPheGlyAlaAlaValGly 80
```

```

QY 318 TCCTTCGCGCGGCGGCGGACGTCGAGTCGCGCGCGGTACACCATCGTGTGCGC 377
Db 81 AlaValGlySerGlyTrpLeuSerPheLysLeuGlyArgLysLysSerLeuMetIleGly 100
QY 378 GCGTCATATCTTCGCGGGGGGTTCCTCATGGGTTCGCGGTCACTACGCGATGTC 437
Db 101 AlaIleLeuPheValAlaGlySerLeuPheSerAlaAlaProAsnValGluIleLeu 120
QY 438 ATGTTTCGCGCGTTCGTCGCGCGCATCGGCTGGGCTACGCGCTCATCATCGCGCGTG 497
Db 121 LeuValSerArgValLeuLeuGlyLeuAlaValGlyValAlaSerTyrThrAlaProLeu 140
QY 498 TACACCGCGAGGTTCGCGCGGTGCGCGGTGCTTCCTGACGCTTCCTCCGAGGTG 557
Db 141 TyrLeuSerGluIleAlaProLysIleArgGlySerMetIleSerMetTyrGlnLeu 160
QY 558 TTCATCAACTTCGCGCATCTGCTCGGTACGTCGAACATATGCTTTCTCCGCTTGCGG 617
Db 161 MetIleThrIleGlyIleLeuGlyAlaTyrLeuSerAspThrAlaPheSer----- 177
QY 618 CTGAACCTCGGCTGCGCATCATGCTCGCATCGCGCGCGCGCTCGTCTGCTGCTCGG 677
Db 178 TyrSerGlyAlaTrpArgTrpMetLeuGlyValIleIleProAlaValLeuLeu 197
QY 678 CTCATGGTCTCGGCATCGCGGAGTCCGCGGTGCTGCTCATGAAGGAGCGCTCGCG 737
Db 198 IleGlyValIlePheLeuProAspSerProArgTrpPheAlaAlaLysArgPheVal 217
QY 738 GACGCCAAGGTGTGCTCGAGAAGACCTCCGACACGCGGAGAGCGCGCGAGCGCTG 797
Db 218 AspAlaGluArgValLeuLeuArgLeuArgAspThrSerAlaGluAlaLysArgGluLeu 237
QY 798 GCGACATCAAGCGCGCGCGCATCTCGAGAGCTCGACGCGCGCTGCTGACCGTC 857
Db 238 AspGluIleArgGluSerLeuLysValLysGln----- 248
QY 858 CCCAAGAGAGGAGCGGAAACGAGAGCGGTGTGGAAGAGGAGCTCATCTGCTCCCGACC 917
Db 249 -----SerGly-----TrpSer-----LeuPheLysAspAsn 257
QY 918 CCGGCATCGCGCGCATCTGCTGCGGATCGGCATCCACTCTTCCACGATCGGTG 977
Db 258 SerAsnPheArgAlaValPheLeuGlyIleLeuLeuGlnValMetGlnGlnPheThr 277
QY 978 GGCATTCTACTCGTGTCTTACAGCCCTCTCGTGTTCAGAGCCCGCGGATTAAACGAAC 1037
Db 278 GlyMetAsnValIleMetTyrTrpAlaProLysIlePheGluLeuAlaGlyTyrAlaAsn 297
QY 1038 GACAAACACTTCTTGGGCACCACTTGGCGGTTCGCTGTCACCAAGAGCGCTTTCATCTTG 1097
Db 298 ThrThrGluGlnMetTrpGlyThrValIleValGlyLeuThrAsnValLeuAlaThrPhe 317
QY 1098 TTGGGACTTCTTTCATCGAGCGGTTCGCGCGCGCGCTGCTGCTGGCAGCAGCGGC 1157
Db 318 IleAlaIleGlyLeuValAspArgTrpGlyArgLysProThrLeu----- 332
QY 1158 GGGATAATCTCTCTCCCTCATCGGCTCGCGCGCGGTTCACCGTGTGTCGCGC-----CAG 1211
Db 333 -----IleLeuGlyPheIleValMetAlaAlaGlyMetGlyValLeuGlyThrMetMet 350
QY 1212 GACCCGAGCCACAGATACCTTGGCGCATCGGCTAAGCATCGCTCCACCTCGCTAC 1271
Db 351 HisIleGlyIleHisSerSerThrAlaGlnTyrIleAlaValLeuMetLeuLeuMetPhe 370
QY 1272 GTCGCTCTTCTTCCATCGGCTTGGCGCCCATCGCGGTGTACAGTCCGAGATCTTC 1331
Db 371 IleValGlyPheAlaMetSerAlaGlyProLeuIleTrpValLeuCysSerGluIleGln 390
QY 1332 CCGCTCCAGGTGCGCGCGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1391
Db 391 ProLeuLysGlyArgAspPheGlyIleThrCysSerThrAlaThrAsnTrpIleAlaAsn 410
QY 1392 GCGCTCATCTCCATGACCTTCTGCTGCTGTCTCAAGGCGCATCACCATCGCGCGCGAGCTTC 1451
```

Db 411 MetIleValGlyAlaThrPheLeuThrMetLeuAsnSerLeuGlySerAlaAsnThrPhe 430
QY 1452 TTCCTCTACTCGGATCCCGCGCTCGCTGGTGTCTTACACCTACCTCCGCGAG 1511
Db 431 TtpValTyrGlyGlyLeuAsnValPheIleLeuLeuThrLeuTrpLeuIleProGlu 450
QY 1512 ACCCGCGCGGACGCTGAGGAGATGACCAAG 1544
Db 451 ThrLysAsnValSerLeuGluHisIleGluArg 461

RESULT 11
US-08-928-692-13
Sequence 13, Application US/08928692
Patent No. 5958727
GENERAL INFORMATION:
APPLICANT: Brody, Howard
APPLICANT: Yaver, Deborah S.
APPLICANT: Lamsa, Michael
APPLICANT: Hansen, Kim
TITLE OF INVENTION: Methods for Modifying the Production of
TITLE OF INVENTION: a Polypeptide
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5958727o No. 5958727disk of No. 5958727th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,692
FILING DATE: 12-SEPT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4944.200-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5958727e
US-08-928-692-13

Alignment Scores:
Pred. No.: 4,23e-36 Length: 584
Score: 576.00 Matches: 148
Percent Similarity: 50.40% Conservative: 106
Best Local Similarity: 29.37% Mismatches: 206
Query Match: 15.94% Indels: 44
DB: 2 Gaps: 11

US-10-051-902A-21 (1-2017) x US-08-928-692-13 (1-584)

QY 153 TTCGCCTTCGCTCGGCGCATCTCGCTCCATGACCTCCATCTCTCGCTACGATATC 212
Db 85 PheIleIleThrLeuThrPheValAlaSerIleSerGlyPheMetPheGlyTyrAspThr 104
QY 213 GGGGTGATGACGGGGCGCTCGCTGATCATCAAGAGGACTTCAC-----ATCAGT 263
Db 105 GlyTyrIleSerSerAlaLeuIleSerIleGlyThrAspLeuAspHisLysValLeuThr 124

QY 264 GACGGGAAGTGGAGGTTCATCGGGCATACTGAACCTCTACTCGCTCATCGCTCCTTC 323
Db 125 TyrGlyGlyGluIleValThrAlaAlaThrSerLeuGlyAlaLeuThrSerIle 144
QY 324 CGCGGGGGCGGACGCTCGGACTCGGCGGGGTACACCATCGTGTTCGCGCGCGTC 383
Db 145 PheAlaGlyThrAlaAlaAspIlePheGlyArgLysArgCysLeuMetGlySerAsnLeu 164
QY 384 ATATTCTTCGCGGGGGTTCCTCATCGGGTTCGCGCTCAACTACGACCATGCTCATGTC 443
Db 165 MetPheValIleGlyAlaIleLeuGlnValSerAlaHisThrPheTrpGlnMetAlaVal 184
QY 444 GCGCGCTTCGCGCGCGCATCGGCTGGCTACCGGCTCATGATCGCGCGCGGTACACC 503
Db 185 GlyArgLeuIleMetGlyPheGlyValGlyIleGlySerLeuIleAlaProLeuPheIle 204
QY 504 GCGAGGTGTCGCGCGGCTCGGCGCTGCTTCCTGACGTCTTCGCGAGGTTCATC 563
Db 205 SerGluIleAlaProLysMetIleArgGlyArgLeuThrValIleAsnSerLeuTrpLeu 224
QY 564 AACTTCGGCATCTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611
Db 225 ThrGlyGlyGlnLeuValAlaTyrGlyCysGlyAlaGlyLeuAsnTyr----- 240
QY 612 TTGCGCTGAACCTCGGCTGCGCATCATGCTCGGCTCGGCGCGCGCGCTCGCTGCTG 671
Db 241 -----ValAsnAsnGlyTrpArgIleLeuValGlyLeuSerLeuIleProThrAlaVal 258
QY 672 CTCGCGCTCATGCTGCTGCGCATCGGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 731
Db 259 GlnPheThrCysLeuCysPheLeuProAspThrProArgTyrTyrValMetLysGlyAsp 278
QY 732 CTCGCGGACGCGGAGGTGCTGCTGAGAGACG---TCGACACGGCGGAGAGCGCGCG 788
Db 279 LeuAlaArgAlaThrGluValLeuLysArgSerTyrThrAspThrSerGluGluIle 298
QY 789 GAGCCCTCGGCGCATCAAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 848
Db 299 GluArgLysValGlu-----GluLeu 305
QY 849 GTGACGCTCCCAAGAGA-----GGAGCGGAAACAGAGAGCGGTGTGAGAGAGTC 902
Db 306 ValThrLeuAsnGlnSerIleProGlyLysAsnValProGluLysValTrpAsnThrIle 325
QY 903 ---ATCCTGTCGCGCGCGCGCATCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
Db 326 LysGluLeuHisThrValProSerAsnLeuArgAlaLeuIleGlyCysGlyLeuGln 345
QY 960 TTCTTCCAGCATGCTGCGCATTCATCCGTGCTTCTACAGCCCTCTCGTGTCAAG 1019
Db 346 AlaIleGlnPheThrGlyTrpAsnSerLeuMetTyrPheSerGlyThrIlePheGlu 365
QY 1020 AGCCCGGATTAACGACGACAAACACTTCTTGGGCACCATTCGCGCTTCGCTGTCACC 1079
Db 366 ThrValGlyPheLysAsnSer-----AlaValSerIleIleValSerGlyThr 382
QY 1080 AAGAGCTTTTCATCTTGTGGCGACTTTCTTCATCGAGCGGCTCGGCGCGCGCGGTG 1139
Db 383 AsnPheIlePheThrLeuValAlaPhePheSerIleAspLysIleGlyArgArgThrIle 402
QY 1140 TTGCTGGGAGCAGCGGCGGGATATCTCTCCCTCATCGGCTCGGCGCGCGGTCCACC 1199
Db 403 LeuLeuIleGlyLeuProGlyMetThrMetAlaLeuValValCysSerIleAlaPheHis 422
QY 1200 GTCGTCGCGCGACCG 1232
Db 423 PheLeuGlyIleLysPheAspGlyAlaValAlaValValSerSerGlyPheSerSer 442
QY 1233 TGGGCGCATCGGCTAAGCATCGCCTCCACCGCTCGGCTCGGCTCTTCTTCTCCATCGCG 1292
Db 443 TrpGlyIle---ValIleIleValPheIleIleValPheAlaAlaPheTyrAlaLeuGly 461
QY 1293 CTGCGCCCATCAGCTGGGTGTACGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352

```

Db 462 ileGlyThrValProTrrp---GlnGlnSerGluLeuPheProGlnAsnValArgGlyIle 480
QY 1353 GGCTGCTCGCTCGCGCGTCCGCGCAACCGCGTCCACCGAGCGGTCATCTCCATGACCTTC 1412
Db 481 GlyThrSerTyAlaThrAlaThrAsnTrpAlaGlySerLeuValIleAlaSerThrPhe 500
QY 1413 CTGTCGCTGTCGAAGGCATACACATCGCGGGGAGCTTCTCTACTCCGCATCGCC 1472
Db 501 LeuThrMetLeuGlnAsnIleThrProAlaGlyThrPheAlaPheAlaGlyLeuSer 520
QY 1473 GCGCTCGCTCGGGTTCCTCTACACCTACCTCCCGGAGACCCGCGCGGACGCTGGAG 1532
Db 521 CysLeuSerThrIlePheCysTyPheCysTyPProGluLeuSerGlyLeuGluLeuGlu 1540
QY 1533 GAGATGAGCAAGCTGTTTCGCGGACACAGCGCGCGCTCGGAATCAGACGAGCCAGCAAG 1592
Db 541 GluValGlnThrIleLeuLysAspGlyPheAsnIleLysAlaSerLysAlaLeuAlaLys 1560
QY 1593 GAGAAGAGGAAG 1604
Db 561 LysArgLysGln 564

RESULT 12
US-09-339-972-13
; Sequence 13, Application US/09339972
; Patent No. 6323002
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 63230020 No. 63230020disk of No. 63230020th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/339,972
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/928,692
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6323002e
US-09-339-972-13

Alignment Scores:
Pred. No.: 4.23e-36 Length: 584
Score: 576.00 Matches: 148

```


Db 363 GlnLeuValValGlyThrLeuIleAlaValLysPheGlyThrSerGlyValGlyGluMet 382
QY 1215 CCCGACGCGCAAGATACCTTGGCGCCATCGGCTAAGCATCGGCTCCACCCCTC----- 1265
Db 383 ProLys-----GlyTyrAlaAlaValValPheIleCys 395
QY 1266 GCCTACGTGGCTTCTTCCATCGGCTGGCCCATCAGCTGGGTGTACAGCTCGGAG 1325
Db 396 LeuTyrValAlaGlyPheAlaTrpSerTrpGlyProLeuGlyTrpLeuValProSerGlu 415
QY 1326 ATCTTCCCGCTCCAGGTGCGGCGCTGGCTGCTCGGCGTCGCGCGCCACCGCGTC 1385
Db 416 IlePheProLeuGluIleArgProAlaGlyGlnSerIleAsnValSerValAsnMetLeu 435
QY 1386 ACCAGCGCGTCATCTCCATGACCTTCTCTGCTGTCCAGGCGCATCACCATCGGCGGC 1445
Db 436 PheThrPheValIleAlaGlnAlaPheLeuThrMetLeuCysHisMetLysPhe--Gly 454
QY 1446 AGCTTCTTCTCTACTCGGCGATCGCGCGCTCGCTCGCTGGCTGTCTTCTACACCTACCTC 1505
Db 455 LeuPheTyrPhePheAlaGlyTrpValValIleMetThrValPheIleAlaLeuPheLeu 474
QY 1506 CCGGAGACCGCGCGCGACGCTCGGAGGAGATG 1538
Db 475 ProGluThrLysAsnValProIleGluGluMet 485

RESULT 14

US-09-679-686B-22

; Sequence 22, Application US/09679686B

; Patent No. 6624343

; GENERAL INFORMATION:

; APPLICANT: Allen, Stephen M.

; APPLICANT: Lightner, Jonathan E.

; APPLICANT: Rafalski, J. Antoni

; APPLICANT: Thorpe, Catherine J.

; FILE OF INVENTION: HEXOSE CARRIER PROTEINS

; TITLE REFERENCE: BB160 US NA

; CURRENT APPLICATION NUMBER: US/09/679,686B

; PRIOR FILING DATE: 2003-01-16

; PRIOR APPLICATION NUMBER: 60/081,131

; PRIOR FILING DATE: 1998-04-09

; PRIOR APPLICATION NUMBER: PCT/US99/07561

; PRIOR FILING DATE: 1999-04-07

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 22

; LENGTH: 514

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-09-679-686B-22

Alignment Scores:

Pred. No.:	4,76e-33	Length:	514
Score:	536.00	Matches:	142
Percent Similarity:	48.79%	Conservative:	99
Best Local Similarity:	28.74%	Mismatches:	193
Query Match:	14.83%	Indels:	60
DB:	4	Gaps:	14

US-10-051-902a-21 (1-2017) x US-09-679-686B-22 (1-514)

QY 159 TTCCCTGGCCATCTCGCTCCATGACCTCATCTCTCGGCTACGATATCGGGTG 218
Db 24 PheValThrCysPheIleGlyAlaPheGlyLeuIlePheGlyTyrAspLeuGlyIle 43
QY 219 ATGAGCGGCGCGCTGGT-----TACATCAAGAG 248
Db 44 SerGlyGlyValThrSerMetGluProPheLeuGluGluPhePheProTyrValTyrLys 63
QY 249 GACTTCAACATCATGT-----GACGGGAAGGTGGAGTTCTC 284
Db 64 LysMetLysSerAlaHisGluAsnGluTyrCysArgPheAspSerGlnLeuThrLeu 83


```
QY 1322 GGAGATCTTCCCGCTCCAGGTCGGCGGTGGCTGCTCGCTCGCGCTCGCGCCACCG 1381
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
416 rGlullePheProLeuGluileArgProAlaGlyGlnSerileAsnValSerValAsnMe 436
QY 1382 CGTCACCGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGG 1441
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
436 tPhePheThrPheCysIleAlaGlnAlaPheLeuThrMetLeuCysHisPheLysPhe-- 455
QY 1442 CGGCAGCTTCTTCCTCTACTCCGGCATCGCGCGCTCGCGCTGGGTGTTCTTACACCTA 1501
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
456 -GlyLeuPheTyrPhePheAlaGlyTrpValValileMetThrValPheIleAlaPhePh 475
QY 1502 COTCCCGGAGACCGCGCGCGGACGCTGGAGGAGATG 1538
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
475 eLeuProGluThrLysAsnValProIleGluGluMet 487
```

Search completed: June 30, 2004, 18:54:21
Job time : 64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:42:00 ; Search time 38 Seconds
(without alignments)
10211.490 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 3614
Sequence: 1 cttacatgtaagctgctgcc.....aaaaaaaaaaaaaaaaaaaaa 2017

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=francet_n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO.spool_P/US10051902/tunat 30062004.164723 20424/app.query.fasta_1.2183
-DB=PIR 78 -QMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCALLCN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051902@cgn 1 1 52 @tunat 30062004.164723 20424 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSFLOCKS=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1589	44.0	549	2	TI4606
2	1579	43.7	511	2	H94536
3	1559	43.1	511	2	A94537
4	1385.5	38.3	508	2	G84564
5	1361	37.7	493	2	A95433
6	1143	31.6	547	2	C84533
7	673.5	18.6	521	2	G84864
8	670.5	18.6	457	2	E70070
9	629.5	17.4	582	2	F71431
10	626	17.3	580	2	D86426
11	609.5	16.9	461	2	D70073
12	609.5	16.9	473	2	G69789
13	601	16.6	557	2	T38125
14	586.5	16.2	580	2	D84772

15	584.5	16.2	472	2	S47089
16	584.5	16.2	560	2	T51485
17	581	16.1	606	2	T27072
18	579.5	16.0	472	2	B26430
19	579.5	16.0	472	2	B91091
20	579.5	16.0	472	2	B85936
21	578.5	16.0	464	2	C91106
22	578.5	16.0	464	2	F85951
23	576	15.9	487	2	B96782
24	576	15.9	584	2	S69555
25	574	15.9	491	2	A26430
26	574	15.9	491	2	F91255
27	574	15.9	491	2	B86096
28	572.5	15.8	464	2	F85079
29	572.5	15.8	612	2	B40538
30	572	15.8	519	2	C87350
31	571	15.8	464	2	AC0877
32	568.5	15.7	471	2	AB0868
33	564	15.6	502	2	B70845
34	562	15.6	639	2	T23658
35	560	15.5	517	2	D96539
36	555.5	15.4	482	2	B98803
37	552	15.3	613	2	T27077
38	543.5	15.0	522	2	S12042
39	543.5	15.0	522	2	B96246
40	542	15.0	490	2	T14545
41	539	14.9	433	2	G96812
42	539	14.9	534	2	S38435
43	538.5	14.9	534	2	S14144
44	536	14.8	514	2	S25009
45	534	14.8	472	2	T35662

ALIGNMENTS

RESULT 1

TI4606
Probable sugar transport protein 205 - beet
C:Species: Beta vulgaris (beet)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C:Accession: TI4606; T14617
R:Chieu, T.J.; Bush, D.R.
A:Title: Molecular cloning, immunochemical localization to the vacuole, and expression in
A:Reference number: Z18131; MUID:96351183; PMID:8742332
A:Accession: TI4606
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-549 <CHI>
A:Cross-references: EMBL:U64902; NID:gl778092; PID:gl778093
A:Accession: T14617
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-541, SVQV' <CH2>
A:Cross-references: EMBL:U64903; NID:gl778094; PID:gl778095
C:Genetics:
A:Note: BvCDNA-205; BvCDNA-397
C:Superfamily: glucose transport protein
C:Keywords: sugar transport

Alignment Scores:
Pred. No.: 3,74e-85 Length: 549
Score: 1589.00 Matches: 326
Percent Similarity: 72.23% Conservative: 59
Best Local Similarity: 61.16% Mismatches: 126
Query Match: 43.97% Indels: 22
DB: 2 Gaps: 5

US-10-051-902A-21 (1-2017) x TI4606 (1-549)

QY 78 TCCTCTGTCACCGAGATGCTTCGCCGCGCTG-----CCGAGGCCGTC 125

DB 10 SerAspProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 29

[illegible]

```

Db      400  PheSerIleGlyAlaGlyProValThrTrpValValCysSerGluIulePheProValArg 419
QY      1341  GTGGCGCGCTGGCGTCTCGCTCGGCGTGGCGGCAACCGCGCTCACACGCGCGTCATC 1400
Db      420  LeuAargAlaGlnGlyAlaSerLeuLeuGlyValMetLeuAsnAargLeuMetSerGlyIleIle 439
QY      1401  TCCATGACCTTCCTGCTGCTGCCAAGGCGATCACCATCGGCGGAGCTTCTTCCTCTCTAC 1460
Db      440  GlyMetThrPheLeuSerLeuSerLysGlyLeuThrIleGlyGlyAlaPheLeuLeuPhe 459
QY      1461  TCGGCATCCCGCGCTCGCTCGGCTGGCTCTCTACACCTACCTCCCGGAGACCCGGCGGC 1520
Db      460  AlaGlyAlaAlaAlaAlaIatrpValPhePhePheThrPheLeuProGluThrAArgGly 479
QY      1521  CGGACGCTGGAGAGATGACCAAGCTGTCGGCGACACGCGCGCTCGGAATCAGAC 1580
Db      480  IleProLeuGluGluMetGluThrLeuPheGlySerTyrThrAlaAsnLysLysAsnAsn 499
QY      1581  GAGCCAGCCAGGAGAGAAAGAGAGTGGA 1610
Db      500  SerMetSerLysAspAsnGluValValAsp 509

RESULT 3
A84537
probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: A84537
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii
M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; T
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.;
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thalian
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84537
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-511 <STC>
A:Cross-references: GB:AE002093; NID:g4678209; PIDN:AAD26955.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g16130
A:Map position: 2
C:Superfamily: glucose transport protein

Alignment Scores:
Pred. No.: 2,08e-83 Length: 511
Score: 1559.00 Matches: 305
Percent Similarity: 75.15% Conservative: 79
Best Local Similarity: 59.69% Mismatches: 123
Query Match: 43.14% Indels: 4
DB: 2 Gaps: 3

```

10

378 GCCGTCAATATCTTCGCGGGGGSGTTCCTCATGGGTTCGCGCAACTACGCACTGC 437
Db GlyPhePhePhePheCysGlyAlaLeuLeuMetGlyPheAlaThrAsnTyrProPheIle 119
438 ATGTTGCGCGCTTCGTGGCGGATCGGCGTGGGCTACGGCTCATGATCGCGCGGTG 497
Db MetValGlyArgPheValAlaGlyIleGlyValGlyTyrAlaMetMetIleAlaProVal 139
498 TACACCGCGAGGTGTCCGCGCTCGCGGTGCTTCCTGAGTGTCTCCCGGAGGTG 557
Db TyrThrThrGluValAlaProAlaSerSerArgGlyPheLeuSerPheProGluIle 159
558 TTCATCACTTCGCGCATCTCTCGGTAGTCTCGAATATGCTTCTCCCGCTTCGCG 617
Db PheIleAsnIleGlyIleLeuLeuGlyTyrValSerAsnTyrPhePheAlaLysLeuPro 179
618 CTGAACCTCGGTCGCGCATCTCTCGGATCGGCGCGGCGGCTCGCTGCTGCTCGG 677
Db GluHisIleGlyTyrArgPheMetLeuGlyIleGlyAlaValProSerValPheLeuAla 199
678 CTCATGTCGTCGCGCATCGCGAGTCCGCGGTGCTGCTCATGAAGGACGCTCGCG 737
Db IleGlyValLeuAlaMetProGluSerProArgTyrPheValMetGlnGlyArgLeuGly 219
738 GACGCCAAGGTGTGCTGGAGAAGACCTCGACACGCGGAGGAGCGCGGAGCGCTG 797
Db AsnAlaPheLysValLeuAspLysThrSerAsnThrLysGluGluAlaIleSerArgLeu 239
798 GCCGACATCAAGCGCGCGCGCATCTCTGAGAGCTCGACGCGGACGTGTGACCGCT 857
Db AsnAspIleLysArgAlaValGlyIleProAspAspMetThrAspValIleValVal 259
858 CCCAAGAGAGGAGCGGAACGAGAAGCGGTGTGGAAGAGCTCACTGCTCCCGGACC 917
Db ProAsnLysLysSerAlaGly--LysGlyValTyrPheLysAspLeuLeuValArgProThr 278
918 CCGGCCATCGCGCATCTCTCGGTGTCGCGATCGGATCCACTCTTCAGCATGGTTG 977
Db ProSerValArgHisIleLeuIleAlaCysLeuGlyIleHisPheSerGlnGlnAlaSer 298
978 GGCATTCACCTCGCTCTCTACAGCCCTCTCGTGTCAAGAGCCCGGATTAACGAAC 1037
Db GlyIleAspAlaValLeuTyrSerProThrIlePheSerArgAlaGlyLeuLysSer 318
1038 GACAAACATCTTTCGCGCACCTTGGCGGTTCGCTGTCTACCAAGAGCTTTTCATCTT 1097
Db LysAsnAspGlnLeuLeuAlaThrValAlaValGlyValValLysThrLeuPheIleVal 338
1098 TTGCGCACTTCTTCATCGCGGTTCGCGGGCGGCGGTGTGTCTGGCGACGCGGC 1157
Db ValGlyThrCysLeuValAspArgPheGlyArgArgAlaLeuLeuLeuThrSerMetGly 358
1158 GGGATAATCTCTCTCCATCGGCTCGCGCGCGGCTCACCGTGTGCGCCAGCACCCC 1217
Db GlyMetPhePheSerLeuThrAlaLeuGlyThrSerLeuThrValIleAspArgAsnPro 378
1218 GACGCCAAGATACCTTCGGCCATCGGCTAAGATCGCTCCACCTCGCGCTACGTCGCG 1277
Db GlyGlnThrLeuLysTyrPheIleGlyLeuAlaValThrThrValMetThrPheValAla 398
1278 TTCTTCTCCATCGGCTTCGCCCATCACGTTGGGTGTACGCTCGGAGATCTTCCCGCTC 1337
Db ThrPheSerLeuGlyAlaGlyProValThrTyrValTyrAlaSerGluIlePheProVal 418
1338 CAGTGTGCGGCTGCGCTCTCTCGCGGTTCGCCCAACCGCTCACCGAGCGGCTC 1397
Db ArgLeuArgAlaGlnGlyAlaSerLeuGlyValMetLeuAsnArgLeuMetSerGlyIle 438
1398 ATCTCCATGACCTTCTCTGCTGTCTCAAGCGCATCACCATCGCGCGGAGCTTCTTCCTC 1457
Db IleGlyMetThrPheLeuSerLeuSerLysGlyLeuThrIleGlyGlyAlaPheLeuLeu 458

1578 TACTCGGCATCGCGCGCTCGCTGGGTGTCTTCTACACCTACCTCCCGAGACCGCG 1517
Db PheAlaGlyValAlaValAlaAlaTyrValPhePhePheThrPheLeuProGluThrArg 478
1518 GCGCGGACGCTCGAGAGATGAGCAAGCTGTTCGGCGACACGCGCGCTCGGATCA 1577
Db GlyValProLeuGluGluIleGluSerLeuPheGlySerTyrSerAlaAsnLysLysAsn 498
1578 GACGAGCCAGCAAG 1610
Db AsnValMetSerLysGlyLysGlnValValAsp 509

RESULT 4
G84564
Probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: G84564
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; NID:20083467; PMID:10617197
A:Accession: G84564
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-508 <STO>
A:Cross-references: GB:AE002093; NID:54218010; PIDN:AAD12218.1; GSPDB:GN00139
A:Gene: At2g18480
A:Map position: 2
C:Superfamily: glucose transport protein

Alignment Scores:
Pred. No.: 2,7e-73 Length: 508
Score: 1385.50 Matches: 271
Percent Similarity: 72.41% Conservative: 78
Best Local Similarity: 56.22% Mismatches: 120
Query Match: 38.34% Indels: 13
DB: 2 Gaps: 3

US-10-051-902A-21 (1-2017) x G84564 (1-508)

150 CGTTCGCTTCGCGCGCTCGCGCATCTCGCTCCATGACCTCCATCTCTCGGTACGAT 209
Db LysPheAlaPheGlyCysAlaIleValAlaSerIleIleSerIleIlePheGlyTyrAsp 39
210 ATCGGGGTGATGAGCGGCGCTGCTGATCATCAAGAGAGACTTCAACATCAGTGACGG 269
Db ThrGlyValMetSerGlyAlaGlnIlePheIleIleArgAspAspLeuLysIleAsnAspThr 59
270 AAGTGTGAGGTTCATCGGCGATACCTGAACTCTACTCTGCTCATCGCTCCTTCCGCGCG 329
Db GlnIleGluValLeuAlaGlyIleLeuAsnLeuCysAlaLeuValGlySerLeuThrAla 79
330 GGGCGGAGCTCGGACTGATCGCGGCGGTACACCATCTGTTGTCGCGCGCTCATATTC 389
Db GlyLysThrSerAspValIleGlyArgArgTyrThrIleAlaLeuSerAlaValIlePhe 99
390 TTCGCGGGGGGTTCCTCATGGGTTCGCCGTCACTACGCGCATGCTCATGTTCCGCGCG 449
Db LeuValGlySerValLeuMetGlyTyrGlyProAsnTyrProValLeuMetValGlyArg 119
450 TTCGTTGGCGGCATCGCGGTGCTAGCGCTCATGATCGCGCGGTGACACCGCGAG 509
Db CysIleAlaGlyValGlyValGlyPheAlaLeuMetIleAlaProValTyrSerAlaGlu 139
510 GTGTCGCGCGGTTCGCGGTTCCTGACGTGCTGTCGCGGTTCGCGGTTCATCAATTC 569
Db IleSerSerAlaSerHisArgGlyPheLeuThrSerLeuProGluLeuCysIleSerLeu 159
570 GGCATCTGCTCGGTGCTGCTGCACTATGCTTCTCCGCTTTCGCGCTGAACCTCGGG 629


```
Db 229 ArgPheGlnAspIleLysAlaAlaGlyIleAspProLysCysValAspValVal 248
QY 852 ACCGTCACAGAGAGGAGGCGGAAACGAGAGCGGCTGCGAGAGCTCATCTCTCC 911
Db 249 LysMetGluGlyLysLysHisGlyGlu--GlyValTrpLysGluLeuLeuArg 267
QY 912 CCGACCCCGGCGATCGCGGATCTCTGTCGCGGATCGGCATCCACTCTCTCCAGCAT 971
Db 268 ProThrProAlaValArgValLeuLeuLeuThrAlaLeuGlyIleHisPheGlnHis 287
QY 972 GCGTTGGCATCTACTCGTCTCTTACAGCCCTCTCGTGTTCAGAGCCCGCGATTA 1031
Db 288 AlaSerGlyIleGluAlaValLeuLeuTrpGlyProArgIlePheLysLysAlaGlyIle 307
QY 1032 --ACGAAACGACAAACACTTCTTGGGCACCACTTGGCGGTGTCACCAAGAGGCTT 1088
Db 308 ThrThrLysAspLysLeuPheLeu--ValThrIleGlyValGlyIleMetLysThrThr 326
QY 1089 TTCATCTGTTGGCACTTCTTTCATCAGCGCTCGGCGGCGCGCTGTTCTGCGGC 1148
Db 327 PheIlePheThrAlaThrLeuLeuLeuAspLysValGlyArgArgLysLeuLeuLeuThr 346
QY 1149 AGCAGCGGCGGATACTCTCTCCCTCATCGGCTCGCGCGCGGCTCACCCTCGCGC 1208
Db 347 SerValGlyGlyMetValIleAlaLeuThrMetLeuGlyPheGlyLeuThr---MetAla 365
QY 1209 CAGCACCCCGAGCAATACCTTGGCGCATCTTAAAGCATCGCCTCCACCCCTCGCC 1268
Db 366 GlnAsnAlaGlyGlyLysLeuAlaThrAlaLeuValLeuSerIleValAlaAlaTyrSer 385
QY 1269 TACGTCGCTTCTCTCCATCGGCTTGGCGCCATCAGCTGGGTGTCACCTCGGATC 1328
Db 386 PheValAlaPhePheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGluVal 405
QY 1329 TTCGCGCTCCAGGTGCGCGCGCTGGTGTCTGCTCGCGCTCGCGCAACCGCTCACC 1388
Db 406 PheProLeuLysLeuArgAlaGlnGlyAlaSerLeuGlyValAlaValAsnArgValMet 425
QY 1389 AGCGGCGTATCTCCATGACCTTCTGCTGCTGTCGCTGTCAGGCGCATCACCATCGCGGCGAGC 1448
Db 426 AsnAlaThrValSerMetSerPheLeuSerLeuThrSerAlaIleThrThrGlyAla 445
QY 1449 TTCCTCTCTACTCGGCATCGCGCGCTCGCTGCTGCTGCTTCTTACACTACCTCCCG 1508
Db 446 PhePheMetPheAlaGlyValAlaValAlaValAlaTrpAsnPhePhePheLeuLeuPro 465
QY 1509 GAGACCCGCGCGGACGCTGGAGGATGAGCAAGCTGTTCTC----- 1550
Db 466 GluThrLysGlyLysSerLeuGluGluIleGluAlaLeuPheGlnArgAspGlyAspLys 485
QY 1551 -----GGCGACGCGCGCGCC 1568
Db 486 ValArgGlyGluAsnGlyAlaAla 493

RESULT 6
C84593
A:Gene: probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 03-Jun-2002
C:Accession: C84593
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84593
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-547 <STO>
A:Cross-references: GB:AE002093; NID:g4454470; PIDN:AAD20917.1; GSPDB:GN00139
C:Genetics:
```

```
A:Gene: At2g20780
A:Map position: 2
C:Superfamily: glucose transport protein

Alignment Scores:
Pred. No.: 3,71e-59 Length: 547
Score: 1143.00 Matches: 234
Percent Similarity: 65.85% Conservative: 88
Best Local Similarity: 47.85% Mismatches: 125
Query Match: 31.63% Indels: 42
DB: Gaps: 5

US-10-051-902A-21 (1-2017) x C84593 (1-547)
QY 150 CGGTTTCGCTTCGCTCGCGCATCTCGCTCCATGACCTCCATCTCTCGGTAC--- 206
Db 53 LysTyrValMetAlaCysAlaPheAlaSerLeuAsnValLeuLeuGlyTyrGly 72
QY 206 ----- 206
Db 73 ArgPheTyrLeuTyrAsnArgIleLeuLeuLeuLeuTyrPheValAspLeuGlnLys 92
QY 207 GATATCGGGGTGATGAGCGGGCGCTGTACATCAAGAAAGGACTTCAACATCAGTGAC 266
Db 93 AspValGlyValMetSerGlyAlaValLeuPheIleGlnAspLeuLysIleThrGlu 112
QY 267 GGAAGGTGGAGTTCTCATGGGATACTGAACCTTACTCTGCTCATCGGCTCTTCGCG 326
Db 113 ValGlnThrGluValLeuIleGlySerLeuSerIleLeuSerLeuPheGlySerLeuAla 132
QY 327 GGGGGCGGAGCTCGGACTGTGATCGCGCGGTACACCATCGTGTTCGCGCGCTCATATA 386
Db 133 GlyGlyArgThrSerAspSerIleGlyArgLysTrpThrMetAlaLeuAlaLeuVal 152
QY 387 TTCCTTCGCGGGGSGTTCCTCATGGGTTCGCGCTCAACTAGCCATGCTCATGTTCGGC 446
Db 153 PheGlnThrGlyAlaAlaValMetAlaValAlaProSerPheGluValLeuMetIleGly 172
QY 447 CGCTTCGCGGGCGATCGCGGTGCTACGCGCTCATGATCGCGGTGCTACACGCGC 506
Db 173 ArgThrLeuAlaGlyIleGlyIleGlyLeuGlyValMetIleAlaProValTyrIleAla 192
QY 507 GAGGTGTCGCGCGCGCTCGCGGTTCCTGCTGCTGCTTCGCGAGGTTCATCAAC 566
Db 193 GluIleSerProThrValAlaArgGlyPhePheThrSerPheProGluIlePheIleAsn 212
QY 567 TTCGCGATCTGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
Db 213 LeuGlyIleLeuLeuGlyTyrValSerAsnTyrAlaPheSerGlyLeuSerValHisIle 232
QY 627 GGGTGGCGCATCTGCTCGCATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 686
Db 233 SerTrpArgIleMetLeuAlaValGlyIleLeuProSerValPheIleGlyPheAlaLeu 252
QY 687 CTCGCGATCGCGGATCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
Db 253 CysValIleProGluSerProArgTrpLeuValMetLysGlyArgValAspSerAlaArg 272
QY 747 GTGGTGTGGAGAGACCTTCGACACGCGGAGGAGCGCGGAGCGCTGCGCGACATC 806
Db 273 GluValLeuMetLysThrAsnGluArgAspAspGluAlaGluGluArgLeuAlaGluIle 292
QY 807 AAGCGCGCGCGCGGATCCCTGAGAGGCTCGACGCGCGAGTGTGTGTGTGTGTGTGTGTGT 866
Db 293 GlnLeuAlaAlaAlaHisThrGlu----- 300
QY 867 GGGAGCGGAAACGAGAGCGG---GTGTGGAGGAGTCTATCTGCTCCGACCGCGCGC 923
Db 301 -----GlySerGluAspArgProValTrpArgGluLeu---LeuSerProSerProVal 317
QY 924 ATGCGCGCATCTGCTGCTGCGGATCGGATCCATCTTCTTCCAGCATGCTGTGGCATT 983
Db 318 ValArgLysMetLeuIleValGlyPheGlyIleGlnCysPheGlnGlnIleThrGlyIle 337
```



```
QY 984 CACTCGCTGCTTCTACAGCCCTCTCGTGTTCACAGAGCCCGGATTAACGACGACAAA 1043
Db 338 AsAlaThrValTyrTyrSerProGluIleLeuLysGluAlaGlyIleGlnAspGluThr 357
QY 1044 CACTTCTTGGGACCACTTGGCGCTTGGGTGTCACCAAGAGGCTTTCTATCTTGTGGG 1103
Db 358 LysLeuLeuAlaAlaThrValAlaValGlyValThrLysThrValPheIleLeuPheAla 377
QY 1104 ACTTCTTCTCATGACGGCTGGCGGCGCGCTGTTCGTGGCAGCAGCGCGGGATA 1163
Db 378 ThrPheLeuIleAspSerValGlyArgLysProLeuLeuTyrValSerThrIleGlyMet 397
QY 1164 ATCTCTCTCATCGGCTCGCGCGGCTCACCGTCTGCGGCGAGCACCCCGAGCGC 1223
Db 398 ThrLeuCysLeuPheCysLeuSerPheThrLeuThrPheLeuGlyGlnGly 414
QY 1224 AAGATACCTTGGCCCATCGGCTTAAGCATCGCTCCACCGCTCGCTTCTTC 1283
Db 415 -----ThrLeuGlyIleThrLeuAlaLeuLeuPheValCysGlyAsnValAlaPhePhe 432
QY 1284 TCCATCGGCTTGGCCCATCAGCTGGGTGTACAGCTCGGAGATCTTCGGCTCCAGGTG 1343
Db 433 SerIleGlyMetGlyProValCysTrpValLeuThrSerGluIlePheProLeuArgLeu 452
QY 1344 CGGCGCTGGGCTGCTCGCTCGCGCTCGCGCCACCGGTACACGCGCGCTCACTCC 1403
Db 453 ArgAlaGlnAlaSerAlaLeuGlyAlaValGlyAsnArgValCysSerGlyLeuValAla 472
QY 1404 ATGACCTTCTGCTGTGTTCAGGCGCATCAACATCGGCGGAGCTTCTTCTACTCC 1463
Db 473 MetSerPheLeuSerValSerArgAlaIleThrValGlyGlyThrPhePheValPheSer 492
QY 1464 GGATCGCGCGCTCGCTCGGCTGCTTCTACACCTACTCCCGGAGACCGCGCGCG 1523
Db 493 LeuValSerAlaLeuSerValIlePheValTyrValLeuValProGluThrSerGlyLys 512
QY 1524 ACCTCGAGGAGATGACGAGCTGTT 1550
Db 513 SerLeuGluGlnIleGluLeuMetPhe 521
RESULT 7
G84864
probable membrane transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84864
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, R.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84864
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-521 <STO>
A;Cross-references: GB:AE002093; NID:g2289003; PIDN:AAB64332.1; GSPDB:GN00139
A;Gene: At2g43330
A;Map Position: 2
C;Superfamily: Glucose transport protein
```

```
Alignment Scores:
Pred. No.: 8.59e-32 Length: 521
Score: 673.50 Matches: 160
Percent Similarity: 51.68% Conservative: 101
Best Local Similarity: 31.68% Mismatches: 177
Query Match: 18.64% Indels: 67
DB: 2 Gaps: 8
```

US-10-051-902A-21 (1-2017) x G84864 (1-521)

```
QY 141 GGCAACGTCGGTTCGGCTTCGGCCATCTCGCTCCATGACCTCCATCCTCCTCCTC 200
Db 27 GlyAsn-----TyrIleLeuGlyLeuThrValThrAlaGlyIleGlyGlyLeuPhe 45
QY 201 GCTACGATATCGGGTGTATGAGGGGGCGTGTGTATACATCAAGAGAGATTCACATC 260
Db 46 GlyTyrAspThrGlyValIleSerGlyAlaLeuLeuTyrIleLysAspAspPheGluVal 65
261 -----AGTGACGGAGAGTG 275
QY 66 ValLysGlnSerSerPheLeuGlnValTyrAsnValSerSerPheThrSerSerLysLeu 85
QY 276 GAGGTTCATCGGGCATACAACTCTACTCGCTCATCGCTCTCTTCGCGCGGGCGG 335
Db 86 GluThrIleValSerMetMetAlaLeuValGlyAlaMetIleGlyAlaAlaGlyGlyTrp 105
QY 336 ACGTCGACTGATCGCGCGGTACACATCTGTTCGCGCGGTATTCCTTCGCG 395
Db 106 IleAsnAspTyrTyrGlyArgLysAlaThrLeuPheAlaAspValValPheAlaAla 125
QY 396 GGGSGTTCCTCATGGGTTCCGCTCAACTACGCCATGCTCATGTCGGCGCTTCGTG 455
Db 126 GlyAlaIleValMetAlaAlaAlaProAspProTyrValLeuIleSerGlyArgLeuLeu 145
QY 456 GCCGCTCGCGCTGGGTACGCTCATGATCGCGCGGTGTATACCGCGAGGTGTGCG 515
Db 146 ValGlyLeuGlyValGlyValAlaSerValThrAlaProValTyrIleAlaGluAlaSer 165
QY 516 CGCGCTCGCGCTGGTCTCTACGCTGTTCGGAGGTGTATCACTTCGGCATC 575
Db 166 ProSerGluValArgGlyGlyLeuValSerThrAsnValLeuMetIleThrGlyGlyGln 185
QY 576 CTGCTCGGTTACGCTCAACTATGCTTCTCCGCTTCCGCTGCAACCTCGGTTGGCG 635
Db 186 PheLeuSerTyrLeuValAsnSerAlaPheThrGlnValProGlyThr-----TrpArg 203
QY 636 ATCATGCTCGCATCGCGCGCGCTGCTGTCTGTCTCGCTCATGCTGTCTCGCATG 695
Db 204 TrpMetLeuGlyValSerGlyValProAlaValIleGlnPheIleLeuMetLeuPheMet 223
QY 696 CCGAGTCGCGCGGTGCTGTATGAAGGAGCGCTCGCGGAGCGCCAGGTGTGTGCTG 755
Db 224 ProGluSerProArgTrpPhePheMetLysAsnArgLysAlaGluAlaIleGlnValLeu 243
QY 756 GAGAGAGCTCCGACACGCGGAGGAGCGCGCGGAGCGCTGCGCGCATCAAGCGCGCC 815
Db 244 AlaArgThrTyrAspIleSer-----ArgLeuGluAsp----- 254
QY 816 GCGGCATCCTCGAGGAGCTCGACGGGACGTGTGTGACCGCTCCCGACAGAGAGGCGGA 875
Db 255 -----GluIleAspHisLeuSerAlaAlaGluGluGluLys 267
QY 876 AACGAGAGCGGTGTGGAAGGAGCTCATCTCCCGACCCCGCGCATCGCGCGCATC 935
Db 268 GlnArgLysArgThrValGlyTyrLeuAspValPheArgSerLysGluLeuArgLeuAla 287
QY 936 CTGCTGTCGGGATCGGCATCACTTCTTCAGCATCGTGTGGGATTCATCCTCGCTGTC 995
Db 288 PheLeuAlaGlyAlaGlyLeuGlnAlaPheGlnPheThrGlyIleAsnThrValMet 307
QY 996 TTCTACAGCCCTCTCGTTCAGAGCCCGGATTAAAGACGACAAACACTTCTTGGGC 1055
Db 308 TyrTyrSerProThrIleValGlnMetAlaGlyPheHisSerAsnGlnLeuAlaLeuPhe 327
QY 1056 ACCACTTGGCGCTTCGGGTGTCCCAAGAGGCTTTTCATCTTGTGGGACTTCTTCATC 1115
Db 328 LeuSerLeuIleValAlaAlaMetAsnAlaAlaGlyThrValValGlyIleTyrPheIle 347
QY 1116 GACGCGCTCGCGCGCGCGCTGTCTGTCGCGAGCAGCGCGGATATCTCTCCCTC 1175
Db 348 AspHisCysGlyArgLysLeuAlaLeuSerSerLeuPheGlyValIleSerLeu 367
1176 ATC-----GGCTCCTCGC 1187
```

```

Db      368 LeuLeuSerValSerPhePheGlnSerGluThrSerSerAspGlyGlyLeuTyr 387
QY      1188 GCCGGGTCCACGGTCGTCGGCCAGCAGCCCGACGCCAAGATACCTTGGGGCATCGGCCTA 1247
Db      388 GlyTrpLeuAlaVal-----LeuGlyLeu 395
QY      1248 AGCATCGCTCCACCTCGCTACGTACGCTTCTTCTCCATCGCGCTTGGCCCATCAGC 1307
Db      396 AlaLeu-----TyrIleValPhePheAlaProGlyMetGlyProValPro 410
QY      1308 TGGGTGACAGCTCGGAGATCTCCCGCTCCAGTGGCGGCTGCGCTCGCTCGCGGC 1367
Db      411 TrpThrValAsnSerGluIleTyrProGlnGlnTyrArgGlyIleCysGlyGlyMetSer 430
QY      1368 GTCGCCGCCAACCGCGTCACAGCGGCGTCATCTCCANGACCTTCCTGTCGTGTCACAG 1427
Db      431 AlaThrValAsnTrpIleSerAsnLeuIleValAlaGlnThrPheLeuThrIleAlaGlu 450
QY      1428 GCCATCACCATCGCGGCGAGCTTCTCTACTCCGCGATCGCGGCTCGCTCGGCTG 1487
Db      451 AlaAlaGlyThrGlyMetThrPheLeuIleLeuAlaGlyIleAlaValLeuAlaIle 470
QY      1488 TTCCTTACACCTACCTCCCGGACCGCGCGCGCGATCGGTGGAGGAGATGACGACGCTG 1547
Db      471 PheValIleValPheValProGluThrGlnGlyLeuThrPheSerGluValGluGlnIle 490
QY      1548 TTCGGCGACACGGCC 1562
Db      491 TrpIysGluArgAla 495

RESULT 8
E70070
metabolite transport protein homolog ywtG - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: E70070
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C:Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, V.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, I.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yamamoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbato, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E70070
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-457 <KUN>
A:Cross-references: GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15600.1; PID:G2636109
A:Experimental source: strain 168
C:Genetics:
A:Gene: ywtG
C:Superfamily: glucose transport protein

Alignment Scores:
Pred. NO.: 1.27e-31 Length: 457
Score: 670.50 Matches: 155
Percent Similarity: 53.29% Conservative: 96
Best Local Similarity: 32.91% Mismatches: 185
Query Match: 18.55% Indels: 35
DB: 2 Gaps: 7

```

US-10-051-902a-21 (1-2017) x E70070 (1-457)

```
Db 332 ThrProAlaAlaSerTrpThrThrValIleCysLeuGlyVal-----Phe 346
Qy 1272 GTCCCTTCCTCCATCGCCCTGGCCCATCAGCTGGGTGTACAGCTCGGAGATCTTC 1331
Db 347 IleValValPheAlaValSerTrpGlyProValValTrpValMetLeuProGluLeuPhe 366
Qy 1332 CCGCTCAGGTGGCGCGCTGGCTGCTCCTCGGTGCTCCGCCCAACCGCTCACCAGC 1391
Db 367 ProLeuHisValArgGlyIleGlyThrGlyValSerThrLeuMetLeuHisValGlyThr 386
Qy 1392 GGGTCTATCTCCATGACTTCTCTGCTGCTCCAAAGGCCATCACCATCGCGCGCAGCTTC 1451
Db 387 LeuIleValSerLeuThrTrpProIleLeuMetGluAlaIleGlyIleSerTrpLeuPhe 406
Qy 1452 TTCTCTACTCCGCATCGCGCGCTGGCTGGTGTCTTCTACACCTACCTCCCGAG 1511
Db 407 LeuIleTrpAlaAlaIleGlyIleMetAlaPheLeuPheValArgPheLysValThrGlu 426
Qy 1512 ACCCGCGCGGCGCTGGAGGAGATGAGCAAG 1544
Db 427 ThrLysGlyArgSerLeuGluIleGluGln 437

RESULT 9.
F71431
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C;Accession: F71431
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirm
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: F71431
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-582 <BEV>
A;Cross-references: GB:297341; NID:G2244991; PID:G2245004
C;Genetics:
A;Map position: 4COP9-4G3845
C;Superfamily: glucose transport protein

Alignment Scores:
Pred. No.: 3.19e-29 Length: 582
Score: 629.50 Matches: 166
Percent Similarity: 46.69% Conservative: 102
Best Local Similarity: 28.92% Mismatches: 206
Query Match: 17.42% Indels: 100
DB: 2 Gaps: 8

US-10-051-902A-21 (1-2017) x F71431 (1-582)
Qy 149 CCGTTCCGCTTCCTCGCGCCATCCTCGCTCCATGACCTCCATCCTCTCGGCTACGA 208
Db 25 ProTyrIleMetArgLeu-AlaLeuSerAlaGlyIleGlyGlyLeuLeuPheGlyTyrAs 44
Qy 209 TATCGGGGTGATGAGGGCGCTGGTGTATCATCAAGAGACTTCAACATCATGTCAGCG 268
Db 44 pThrGlyValIleSerGlyAlaLeuLeuPheIleLysGluAspPheAspGluValAspLys 64
Qy 269 GAAGGTGAGGTCTCATGGGTACTGAACCTCTAC-----TCGCTCATCGGCTC 319
Db 64 sLysThrTrpLeuGlnSerThrIleValSerMetAlaValAlaGlyValIleValGlyAl 84
Qy 320 CTTCCGCGCGCGGCGGACGTGGATCGGCGCGGTCACACCATCGTGTTCGCCGC 379
Db 84 aAlaValGlyGlyTrpIleAsnAspLysPheGlyArgArgMetSerIleLeuIleAlaAs 104
```

```
Qy 380 CQTCTATTCTTCGCGGGGGSTTCTCTCATGGGTTCGCCCTCAACTACGCCATCTCAT 439
Db 104 pValLeuPheLeuIleGlyAlaIleValMetAlaPheAlaProAlaProTrpValIle1 124
Qy 440 GTTCGGCGCTTCGTCGGCGGCATCGGCTGGGTACCGCTCATGATCGCGCGGTGTA 499
Db 124 eValGlyArgIlePheValGlyPheGlyValGlyMetAlaSerMetThrSerProLeuTy 144
Qy 500 CACCCCGCAGGTGTCCGCGGCTCGCGCTGGCTTCCTGACCTGCTTCCCGGAGGTGT 559
Db 144 rIleSerGluAlaSerProAlaArgIleArgGlyAlaLeuValSerThrAsnGlyLeuLe 164
Qy 560 CATCAACTTCGGCATCTCGGTGTCTCGAGCTATGCTTCTCCGCTTCGCCGT 619
Db 164 uIleThrGlyGlyGlnPhePheSerTrpLeuIleAsnLeuAlaPheValHisThrProG 184
Qy 620 GAACCTCGGGTGGCGCATCATGCTCGCATCGCGCGCGCGCTCGCTCGCTCGCT 679
Db 184 yThr-----TrpArgTrpMetLeuGlyValAlaGlyValProAlaIleValGlnPheVa 202
Qy 680 CATGTGTTCGGCATCGCGGAGTCCCGGTGTGTGTATGAAGGACGCTCGCGGA 739
Db 202 lLeuMetLeuSerLeuProGluSerProArgTrpLeuTrpArgLysAspArgIleAlaG 222
Qy 740 CGCCAAGTGTGTGTGGAGAAG-----ACCTCCGACACGCGGAGGAGCGCGGAGCG 793
Db 222 uSerArgAlaIleLeuGluArgIleTrpProAlaAspGluValGluAlaGluMetGluAl 242
Qy 794 CTTGCGCCGACATCAAGCGCGCGCATCTCTCGGTATCGGAGCTCGACGCGCGCTGTCAC 853
Db 242 aLeuLysLeuSerValGluAlaGluLysAlaAspGluAlaIleIleGlyAspSerPheSe 262
Qy 854 CTTCCCGCAGAGAGGAGC---GGNAACGACGAGCGGTGTGAAGGAGCTCATCTCTC 910
Db 262 rAlaLysLeuLysGlyAlaPheGlyAsn----- 271
Qy 911 CCGGACCGCGGCATCGCGCATCTCTCTCGGTATCGGATCGCATCTCTTCCAGCA 970
Db 272 -----ProValValArgArgGlyLeuAlaGlyIleThrValGlnValAlaGlnGl 289
Qy 971 TCGGTGGGATCTACTCCGCTCTCTTACAGCCCTCTCGTGTTCAGAGCCCGGATT 1030
Db 289 nPheValGlyIleAsnThrValMetTrpTrpSerProSerIleValGlnPheAlaGlyTy 309
Qy 1031 AACGAACGACAAACACTTCTTGGGCACCACTTGGCGCTCGGTGTCCACCAAGAGCTTT 1090
Db 309 rAlaSerAsnLysThrAlaMetAlaLeuSerLeuIleThrSerGlyLeuAsnAlaLeuGl 329
Qy 1091 CATCTTGTGGGACTTTCTTCATCGACGGCTCGCGCGCGCGCTGTGTCTGGGCGAG 1150
Db 329 ySerIleValSerMetMetPheValAspArgTrpGlyArgArgLysLeuMetIleIleSe 349
Qy 1151 CACGGCGGGATATCTCTCTCATCGCTCGCGCGCGGCTCACCGCTCGTCGGCCA 1210
Db 349 rMetPheGlyIleIleAlaCysLeuIleIleLeuAlaThrValPheSerGlnAlaAlaI 369
Qy 1211 GCAC-----CCGACGCGCAAGATA-- 1229
Db 369 eHisAlaProLysIleAspAlaPheGluSerArgThrPheAlaProAsnAlaThrCysSe 389
Qy 1229 ----- 1229
Db 389 rAlaTrpAlaProLeuAlaAlaGluAsnAlaProProSerArgTrpAsnCysMetLysCy 409
Qy 1230 -----CCTTGGGCGCATCGG-- 1244
Db 409 sLeuArgSerGluCysGlyPheCysAlaSerGlyValGlnProTrpAlaProGlyAlaCy 429
Qy 1244 ----- 1244
Db 429 sValValLeuSerAspMetLysAlaThrCysSerSerArgGlyArgThrPhePheLy 449
```



```
Db 407 rProSerCysGlyTyrCysSerSerProIleGlyLysGluHisProGlyAlaCysTrpIl 427
QY 1223 -----
Db 427 eSerAspAspSerValLysAspLeuCysHisAsnGluAsnArgLeuTrpTyrThrArgGl 447
QY 1224 -----AAGATACCTTGG-----CCCATCGCCTTAAGCATCGCCTCCACCTCGC 1267
Db 447 yCysProSerAsnPhaGlyTrpPheAlaLeuLeuGlyLeuLeu----- 462
QY 1268 CTACGTCGGCTCTTCTCCATCGCCTTGGCCCATCAGTCGGGTGTACAGCTCGGAGAT 1327
Db 463 -TyrIleIlePhePheSerProGlyMetGlyThrValProTyrIleValAsnSerGluIl 482
QY 1328 CTTCGGCTCCAGTCGGCGCTGGCTGCTCGCTCGCGCTGGCGCGCAACCGGTAC 1387
Db 482 eTyProLeuArgPheArgGlyIleCysGlyGlyAlaAlaAlaThrAlaAsnTrpIleSe 502
QY 1388 CAGCGCGTCATCTCCATCGCTTCTCGCTGTCCTCAAGGCATCACCATCGGGGCGAG 1447
Db 502 rAsnLeuIleValAlaGlnSerPheLeuSerLeuThrGluAlaIleGlyThrSerTrpTh 522
QY 1448 CTCTCTCTCTACTCCGGATCGCCGCTCGCTGGGTGTCTTCTACACTACTCTCC 1507
Db 522 rPheLeuIlePheGlyValIleSerValIleAlaLeuLeuPheValMetValCysValPr 542
QY 1508 GGAGACCCGCGCGACGCTGGAGGAGATGACAGCTG 1547
Db 542 oGluThrLysGlyMetProMetGluIleGluLysMet 555

RESULT 11
D70073
metabolite transport protein homolog yxC - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: D70073
A:R;Kunet, S.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bester
A:; Brøn, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A:; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
A:; Kocher, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
A:; Ketter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yamamoto, K.; Yata, K.; Yoshida, K
A:; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D70073
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-461 <UN>
A:Cross-references: GB:Z99124; GB:AL009126; NID:G2636442; PIDN:CA316017.1; PID:G2636527
A:Experimental source: strain 168
C:Genetics:
A:Gene: yxC
C:Superfamily: glucose transport protein

Alignment Scores:
Pred. No.: 4,57e-28 Length: 461
Score: 609.50 Matches: 149
Percent Similarity: 49.59% Conservative: 93
Best Local Similarity: 30.53% Mismatches: 195
Query Match: 16.86% Indels: 51
DB: 2 Gaps: 11

US-10-051-902A-21 (1-2017) x D70073 (1-461)
QY 180 TCATGACCTCCATCTCTCTCGCTACGATATCGGGGTGATGAGCGGGGCGTGGTGTAC 239
```

```
Db 15 AlaLeuGlyGlyLeuLeuTyrGlyTyrAspThrGlyValIleSerGlyAlaLeuPhe 34
QY 240 ATCAAGAGAGCTTCAACATCAGTACGGGAAGTGGAGTTCTCATGGGCATCTGAAC 299
Db 35 IleAsnAsnAspIleProLeuThrThrLeuThrGluGlyLeuValValSerMetLeuLeu 54
QY 300 CTCTACTCGCTCATCGCTCTTCCGCGGGGGGAGCTCGGACTGGATCGCGCGCGG 359
Db 55 LeuGlyAlaIlePheGlySerAlaLeuSerGlyThrCysSerAspArgTrpGlyArgArg 74
QY 360 TACACCATCGTGTTCGCGCGCGCTCATATTCTTCGCGGGGGGGTTCCTCATGGGGTTCGCC 419
Db 75 LysValValPheValLeuSerIleIlePheIleIleGlyAlaLeuAlaCysAlaPheSer 94
QY 420 GTCACATACGCATGCTCATGTTCCGCGCTTCTCGTGGCGGCATCGCGGTGGGTAGCG 479
Db 95 GlnThrIleGlyMetLeuIleAlaSerArgValIleLeuGlyLeuAlaValGlyGlySer 114
QY 480 CTCATGATCGCGCGGTGTACACCGCGAGGTGTCCGCGCGCTCGCGCGCTGGTTCCTG 539
Db 115 ThrAlaLeuValProValTyrLeuSerGluMetAlaProThrLysIleArgGlyThrLeu 134
QY 540 AGTCGCTTCGCGAGGTGTTTCATCACTTCGGCATCTCTGTCGGGTAGCTCTCGAATAT 599
Db 135 GlyThrMetAsnAsnLeuMetIleValThrGlyIleLeuLeuAlaTyrIleValAsnTyr 154
QY 600 GCTTTCCTCCGCTTTCGCTGAACCTCGGGTGGGCATGCTCGGCATCGCGCGCGCG 659
Db 155 LeuPheThr-----ProPheGlu---AlaTrpArgTrpMetValGlyLeuAlaVal 171
QY 660 CCGTCCGTGCTGCTCGCGCTCATGTCGTCCGATGCGGAGTCCGCGGTGGTGGTGGT 719
Db 172 ProAlaValLeuLeuLeuIleGlyIleAlaPheMetProGluSerProArgTrpLeuVal 191
QY 720 ATGAAGGAGACCTTCGCGGAGCCCAAGGTGTGTGGAGAAGACCTCCGACACGCGGAG 779
Db 192 LysArgLysGluGluGluAlaArgGlyMetAsnIleThrHisAsp----- 208
QY 780 GAGCGCGGAGCGCTGGCGGACATCAAGCGCGCGCGCATCTCCCTGAGAGTGCAC 839
Db 209 -----ProLysAspIleGlu 213
QY 840 GCGACGCTGTGACGCTCCCAAGAGAGGAGCGGAAACGAGAAGCGGTGTGAAGGAG 899
Db 214 MetGluLeuAlaGluMet---LysGlnGlyGluAlaGluLys-----LysGlu 228
QY 900 CTCATCTGTCCTCCGACCGCGCC-----ATCGCGGCATCTCTGTCGGGTAGCT 953
Db 229 ThrThrLeuGlyValLeuLysAlaLysTrpIleArgProMetLeuLeuIleGlyValGly 248
QY 954 ATCCACTTCTTCCAGCATGCTGGGCATTCACCTCGTGTCTTCTACAGCCCTCTGCT 1013
Db 249 LeuAlaIlePheGlnAlaValGlyIleAsnThrValIleTyrAlaProThrIle 268
QY 1014 TTCAGAGCCCGGATTAACAAACGACAAACACTTCTTGGGCACCACTTGGCGGTTCG 1073
Db 269 PheThrLysAlaGlyLeuGlyThrSerAlaSerAlaLeuGly---ThrMetGlyIleGly 287
QY 1074 GTCACCAAGAGCGCTTTCATCTTGTGGCGACTTCTTCATCGAGCGGTGGCGCGG 1133
Db 288 IleLeuAsnValIleMetCysIleThrAlaMetIleLeuIleAspArgValGlyArgLys 307
QY 1134 CCGTGTGTGGGAGACGCGGGGGGATATCTCTCCCTCATCGCGCTCGCGCGCGG 1193
Db 308 LysLeuLeuIleTrpGlySerValGlyIleThrLeuSerLeuAlaLeuSerGlyVal 327
QY 1194 CTCACCGCTGTCGCGCAGCAGCCCGAGCAAGATACCTTGGGCATCGCGCTAAGCATC 1253
Db 328 LeuLeuThrLeuGlyLeu-----SerAlaSerThrAlaTrp-----MetThrVal 342
QY 1254 GCCTCCACCTCGCTACGTCGCTCTCTCTCCATCGCGCTTGGCCCATCACCTGGGTG 1313
```

```

Db      343 ValPheLeuGlyValTyrIleValPheTyrClnAlaThrTrpGlyProValValTyrVal 362
QY      1314 TACAGCTCGGAGATCTCCCGCTCCAGTGGCGCGCTGGCTGCTCGCTGGCGGTGCGCC 1373
Db      363 LeuMetProGluLeuPheProSerLysAlaArgGlyAlaAlaThrGlyPheThrThrLeu 382
QY      1374 GCCAACCGCGTCACAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGCCGTCCAAAGGCCATC 1433
Db      383 ValLeuSerAlaAlaAsnLeuIleValSerIleValPheProLeuMetLeuSerAlaMet 402
QY      1434 ACCATCGCGCGACCTTCTCTCTACTCCGCGATCCCGCGCTCGCTGGGTGTTCTTC 1493
Db      403 GlyIleAlaTrpValPheMetValPheSerValIleCysLeuLeuSerPhePheAla 422
QY      1494 TACACCTACCTCCGCGAGACCGCGCGCGACGCTGGAGGAG-----ATCAGC 1541
Db      423 PheTyrMetValProGluThrLysGlyLysSerLeuGluLeuGluAlaSerLeuLys 442
QY      1542 AAGCTGTTTCGCGCAGACGCGCGCGCTCGGAATCAGACGAGCCAGCAAGAGAGAAG 1601
Db      443 LysArgPhe-----LysLysLysLys 449
QY      1602 AAGTGAAATGGCGGCACCTAAC 1625
Db      450 SerThrGlnAsnGlnValLeuAsn 457

RESULT 12
G69789
sugar transporter homolog ydJK - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: G69789
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteux
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaite, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: G69789
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-473 <KUN>
A:Cross-references: GB:Z99107; GB:AL009126; NID:G2632866; PIDN:CAB12442.1; PID:G2632936
A:Experimental source: strain 168
C:Genetics:
A:Gene: ydJK
C:Superfamily: glucose transport protein

Alignment Scores:
Pred. No.: 4,59e-28 Length: 473
Score: 609.50 Matches: 157
Percent Similarity: 52.48% Conservative: 97
Best Local Similarity: 32.44% Mismatches: 187
Query Match: 16.86% Indels: 43
DB: 2 Gaps: 11

US-10-051-902a-21 (1-2017) x G69789 (1-473)
QY      135 AAGAGGGCAAC---GTCGGGTTCGCTTCGCTCGCGCCATCCTCGCTCCATGACCTCC 191
Db      3 LysGlnGlyAsnGlnMetSerPheLeuAlaThrIleIleLeuValSerThrPheGlyGly 22
QY      192 ATCTCTCTCGGTACGATATCGGGGTGATGAGCGGGCGTTCGCTGTACATCAAGAAG--- 248

```

```

Db      23 LeuLeuPheGlyTyrAspThrGlyValLeuAsnGlyAlaLeuProTyrMetGlyGluPro 42
QY      249 ---GACTTCACATCAGTACGAGGAGGTGAGGTTCTCATGGGCATACCTGAACTCTAC 305
Db      43 AspGlnLeuAsnLeuAsnAlaPheThrGluGlyLeuValThrSerSerLeuPheGly 62
QY      306 TCGCTCATCGCTCTCTCGCGCGGGGCGACGTCGAGTCGATCGGCGCGCGGTACACC 365
Db      63 AlaAlaLeuGlyAlaValPheGlyArgMetSerAspPheAsnGlyArgArgLysAsn 82
QY      366 ATCGTGTTCGCGCGCTCATATTCTTCGCGGGGGGTTCCTCATGGGTTCGCGCTCAAC 425
Db      83 IleLeuPheLeuAlaValIlePhePheIleSerThrIleGlyCysThrPheAlaProAsn 102
QY      426 TACGCCATGCTCATGTTTCGGCGCTTCGTCGCGGCATCGGCTGGGTGGGTACGCGCTCATG 485
Db      103 ValThrValMetIleIleSerArgPheValLeuGlyIleAlaValGlyAlaSerVal 122
QY      486 ATCGCGCGCGGTACACCGCGAGGTGTCGCGCGCGTGGCGCGTGGCTTCTCTGACGTG 545
Db      123 ThrValProAlaTyrLeuAlaGluMetSerProValGluSerArgGlyArgMetValThr 142
QY      546 TTCCCGGAGGTGTTCACTCAACTTCGGCATCTGCTCGGGTACGTCGAACTATGCTTTC 605
Db      143 GlnAsnGluLeuMetIleValSerGlyGlnLeuLeuAlaPheValPheAsn----- 159
QY      606 TCCGCTTCGCGCTGAACCTCGGG-----TGGCGCATCATCTCGGCATC 650
Db      160 AlaIleLeuGlyThrThrMetGlyAspAsnSerHisValTrpArgPheMetLeuValIle 179
QY      651 GCGCGCGCGCTCGCTCGCTCGCTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 710
Db      180 AlaSerLeuProAlaLeuPhePhePheMetIleArgMetProGluSerProArg 199
QY      711 TGGCTGTGTCATGAGGAGCGCTCGCGACGCCAAGGTGGTGTGGAGAGACCTCCGAC 770
Db      200 TrpLeuValSerLysGlyArgGlyGluAspAlaLeuArgValLeuLysLysIleArgAsp 219
QY      771 ACGGCGGAGGAGCGCGCGCGCTCGCGCATCAAGCGCGCGCGCGCGCATCCCTCGAG 830
Db      220 GluLysArgAlaAlaAlaGlu---LeuGlnGluIleGluPheAlaPheLysLysGluAsp 238
QY      831 GAGCTCAGCGCGCGCGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 890
Db      239 GlnLeu-----GluLysAlaThr 244
QY      891 TGAAGGAGCTCATCTCTGTCGCCGACCGCGCATCGCGCGCATCTCTGTCGTCGTCGTCG 950
Db      245 PheLysAspLeu-----SerValProTrpValArgIleValPheIleGlyLeu 261
QY      951 GGCATTCATCTTCTCCAGCATCGCTTGGCATTCACCTCGCTGCTGCTCTCTACAGCCCTCTC 1010
Db      262 GlyIleAlaIleValGlnGlnIleThrGlyValAsnSerIleMetTyrTyrGlyThrGlu 281
QY      1011 GTGTTCAAGAGCCCGCGATTAAACGAACCAACACTTCTTGGGCACCACTTGGCGCGTC 1070
Db      282 IleLeuArgAsnSerGlyPheGlnThrGluAlaAlaLeuIleGly----- 296
QY      1071 GGTGTCAACCAAGAGCTTTCATCTTGTGGCGACTTTC-----TTCATCGAC 1118
Db      297 AsnIleAlaAsnGlyValIleSerValLeuAlaThrPheValGlyIleTrpLeuGly 316
QY      1119 GCGCTCGCGCGCGCGCTGTCGTCGCGAGCAGCGCGGAGATAATCTCTCTCTCTCTC 1178
Db      317 ArgValGlyArgAspProMetLeuMetThrGlyLeuIleGlyThrThrAlaLeuLeu 336
QY      1179 GGCCTCGCGCGCGCGCTCACCGCTCGTCGGCGAGCACCCCGAGCCCAAGATCACTTGGGCC 1238
Db      337 LeuIleGlyIlePheSerLeuValLeuGluGlySerPro-----AlaLeuProTyrVal 354
QY      1239 ATCGGCTAAGCATCGCTCCACCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1298

```

Db 355 Val-----LeuSerLeuThrValThrPheLeuAlaPheGlnGlnGlyAlaIleSer 371
QY 1299 CCATACAGTGGGTACAGTCTCGAGATCTTCCCGCTCCAGGTGCGCGCTGGGCTGC 1358
Db 372 ProValThrTrpLeuMetLeuSerGluIlePheProLeuArgLeuArgGlyLeuGlyMet 391
QY 1359 TCCTCGGCGTCCCGCCAAACCGCTCACAGCGGCGTCTATCCATGACCTTCTCTGTCG 1418
Db 392 GlyValThrValPheCysLeuTrpMetValAsnPheAlaValSerPheThrPheProIle 411
QY 1419 CTCTCAAGGCCATCACTACGCGCGAGCTTCTCTCTACTCCGCGCATCGCGCGCTC 1478
Db 412 LeuLeuAlaAlaIleGlyLeuSerThrThrPhePheIlePheValGlyLeuGlyIleCys 431
QY 1479 GCTGGGTGTTCTTACACTTCCCGGAGACCGCGCGGCGGAGCTGAGGAGATG 1538
Db 432 SerValLeuPheValLysArgPheLeuProGluThrLysGlyLeuSerLeuGluGlnLeu 451
QY 1539 AGCAAGCTGTC 1550
Db 452 GluGluAsnPhe 455

RESULT 13
T38125
myo-inositol transporter 2 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, May 1997
A:Reference number: Z21772
A:Accession: T38125
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-557 <BAD>
A:Cross-references: EMBL:Z95334; PIDN:CA08597.1; GSPDB:GN00066; SPDB:SPAC20G8.03
A:Experimental source: strain 972h-; cosmid c20G8
C:Genetics:
A:Gene: SPDB:SPAC20G8.03
A:Map position: 1
C:Superfamily: maltose transport protein MAL61

Alignment Scores:
Pred. No.: 1.46e-27 Length: 557
Score: 601.00 Matches: 160
Percent Similarity: 51.22% Conservative: 112
Best Local Similarity: 30.13% Mismatches: 211
Query Match: 16.63% Indels: 48
DB: 2 Gaps: 16

US-10-051-902A-21 (1-2017) x T38125 (1-557)
QY 93 GAGATGCTTCGCGCGCTGCGGAGCGCTGCGCGCAAGAGGCAACGTCCGG 152
Db 52 GluGluAsnAlaThrAspProGlnAlaAsnGluValAlaAspGluAsnGlyGluGly 71
QY 153 TTC-----GCCTTCGCTCGCCATCTCTCGCTCCATGAC 188
Db 72 PheGluAlaGluLysIleSerSerTrpIleTrpValLeuSerAlaValAlaGlyIleSer 91
QY 189 TCCATCTCTCGCTACGATATCGGCTGATGAGCGGCGCTGCTGTACATCAAGAG 248
Db 92 GlyLeuLeuPheGlyTyrAspThrGlyValIleSerGlyAlaLeuAlaValLeuGlySer 111
QY 249 GACTTCAAC-----ATCAGTACGGAAGGTGAGGTCTCTCATGGGCATACCTGAACTC 302
Db 112 AspLeuGlyHisValLeuSerSerGlyGlnLysGluLeuIleThrSerAlaThrSerPhe 131
QY 303 TACTGCTATCGCTCTCTCGCGCGGGGAGCTGCGACTGATCGCGCGGCTAC 362
Db 132 AlaAlaLeuIleSerAlaThrThrSerGlyTrpLeuAlaAspTrpValGlyArgLysArg 151
QY 363 ACCATCGTGTTCGCGCGCTCATATTCTTCGCGGGGGTTCCTCATGGGTTCGCGCTC 422

Db 152 LeuLeuLeuCysAlaAspAlaIlePheValIleGlySerValIleMetAlaAlaSerArg 171
QY 423 AACTACGCGCATGCTCATGTTTCGCGCGCTTCTGCGCGGATCGCGGTGGGTACGCGCTC 482
Db 172 AsnValAlaMetMetValValGlyArgPheIleValGlyTyrGlyIleGlyLeuThrSer 191
QY 483 ATGATCGCGCGGTGTACACCGCGAGGTGTCGCGCGCTGCGCGGTGCTCTGACG 542
Db 192 LeuIleValProMetTyrIleThrGluLeuAlaProAlaArgLeuArgGlyValLeuVal 211
QY 543 TCGTTCGCGGAGTGTTCATCAACTTCGGAATCTGTCGCGGTACGTCGAACTATGCT 602
Db 212 IleIleTyrValValPheIleThrGlyGlyLeuIleAlaTyrSerLeuAsnAlaAla 231
QY 603 TTCCTCCGCTTGGCGCTGAACCTCGGTGGCGCATGCTCGGCATCGCGCGCGCGCG 662
Db 232 PheGluHis-----ValHisGlnGlyTrpArgIleMetPheGlyIleGlyAlaAlaPro 249
QY 663 TCCGTGCTGCTCGCGCTCATGTCGCGCATGCGGAGTTCGCGCGGTGCTGTCATG 722
Db 250 AlaLeuGlyGlnLeuIleSerLeuPheTrpThrProGluSerProArgTyrLeuLeuArg 269
QY 723 AAGGACCCCTCGCGGACGCCCAAGTGTGTGTGAGAGAGACTCGACACGCGGAG--- 779
Db 270 HisAsnHisValGluLysValTyrLysIleLeuSerArgIleHisProGluAlaLysPro 289
QY 780 ---GAGCGCGGAGCGCTGCGCGACATCAAGGCGCGCGCGCATCCTCGAGGAGCTC 836
Db 290 AlaGluIleAlaTyrLysValSerLeuIleGln----- 300
QY 837 GACGCGAGCTGTGACCGTCCCAAGAGAGGAGCGGAAACAGAG---AAGGGGTGTGG 893
Db 301 GluGlyValLysValAspPheProGlu-----GlyAsnLysPheHisPhePhe 317
QY 894 AAGGAGCTC---ATCCTCTCCCGACCCCGCGCATGCGCGCATCCTCTGTCGCGGATC 950
Db 318 HisSerLeuLysValLeuPheThrValProSerAsnArgSerLeuPheIleGlyCys 337
QY 951 GGCATCCACTTCTTCCAGCATGCGTGGCGATTCACCTCGTCTGCTTCTACAGCCCTCTC 1010
Db 338 PheLeuGlnTrpPheGlnGlnPheSerGlyThrAsnAlaIleGlnTyrPheSerAlaIle 357
QY 1011 GTGTTCAGAGCCCGGATTAACGAAACGACAAACACTTCTTGGGACACACTTGGCGGTTC 1070
Db 358 IlePheGlnSerValGlyPheLysAsnSer-----IleSerValSerIleValVal 374
QY 1071 GGTGTCCACAGAGCTTTTCATCTGTGTGGCGACTTCTTTCATCGAGCGGTTCGCGGG 1130
Db 375 GlyAlaThrAsnPheValPheThrIleValAlaPheMetPheIleAspArgIleGlyArg 394
QY 1131 CGCGCGCTGTGCTGGGAGACGCGGCGGATATCTCTCTCCCTCATCGGCTCGCGCGCC 1190
Db 395 ArgArgIleLeuLeu---CysThrSerAlaValMetIleAla-----GlyLeu 409
QY 1191 GGGCTCACGCTGCTCGCGCAGCAC-----CCGACGCCCAAGATA 1229
Db 410 AlaLeuCysAlaIleAlaTyrHisPheLeuProAlaAspThrThrGlnAsnThrAsnSer 429
QY 1230 CCTTGGGCGCATCGGCTAAGCATCGCTCCACCCCTCGCTACGTCGCTCTCTCTCCATC 1289
Db 430 GlyTrpGlnTyr---ValValLeuAlaSerIleIleIlePheLeuAlaSerTyrAlaSer 448
QY 1290 GGCCTTGGCGCCCATCAGTGGGTGTACAGCTCGGAGATCTTCCGCTCCAGTGGCGCG 1349
Db 449 GlyIleGlyAsnIleProTrp---GlnGlnAlaGluLeuPheProMetGluValArgAla 467
QY 1350 CTGGGTGCTGCTCGGCGTCCCGCCCAACCGCTCACGCGGTTCATCTCCATGACC 1409
Db 468 LeuGlyAlaGlyPheSerThrAlaIleAsnTrpValGlyAsnLeuIleIleSerAlaSer 487
QY 1410 TTCCTGTGCTGTCAAGGCCCATCAGTCGCGCGGAGCTTCTCTCTACTCCGCGATC 1469

Db 488 PheLeuThrMetMetGluSerIleThrProThrGlyThrPheAlaLeuPheAlaGlyPhe 507
 QY 1470 GCCGGCTCGCTGGTGTCTTCTACACCTACCTCCGGAGACCCGGCGGACGCTG 1529
 Db 508 CysPheValGlyLeuValThrSerTyrPheThrTyrProGluLeuAlaGlyMetSerIle 527
 QY 1530 GAGGAGATGACCAAGCTGTC-----GGGACACGGCGCGCTCGGAATCAGACGAG 1583
 Db 528 GluAsnIleHisLysLeuLeuGluLysGlyPheThrPheGlnAlaValLysGluSerThrLys 547
 QY 1584 CCAGCCCAAGGAGAGAAAGAGTGGAAATGGCC 1616
 Db 548 ArgValArg--LysGlyArgIleAspGluAla 557
 RESULT 14
 Db4772
 probable sugar transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
 C:Accession: D84772
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 euss, D.; Mierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; PMID:20083487; PMID:10617197
 A:Accession: D84772
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-580 <STO>
 A:Cross-references: GB:AB002093; NID:94263781; PIDN:AAD15441.1; GSPDB:GN00139
 C:Gene: At2g35740
 A:Map position: 2
 C:Superfamily: glucose transport protein
 Alignment Scores:
 Pred. No.: 1.02e-26 Length: 580
 Score: 586.50 Matches: 154
 Percent Similarity: 46.38% Conservative: 102
 Best Local Similarity: 27.90% Mismatches: 198
 Query Match: 16.23% Gaps: 98
 DB: 2
 US-10-051-902a-21 (1-2017) x D84772 (1-580)
 QY 143 CAACGTCGGTTCGCTCGCTCGCGCCATCTCGCTCCATCGCTCCATCTCCCTCGG 202
 Db 22 GluThrProTyrIleMetArgLeu-AlaLeuSerAlaGlyIleGlyLeuLeuPheG1 41
 QY 203 CTACATATCGGGTGTATGACGGGGCGTCTGTCTATCATCAAGAGACTTCAACATCAG 262
 Db 41 YTYrAsnThrGlyValIleAlaGlyAlaLeuLeuTyrIleLysGluLupheGlyLuVa 61
 QY 263 TGACGGGAAGTG-----GAGTTCATCGGGGCTACTGACCTCTACTCGTCTCAT 313
 Db 61 LAsPasnLysThrTrpLeuGlnGluIleValSerMetThrValAlaGlyAlaIleVa 81
 QY 314 CGGTCCTTCGCGGGGGGAGCTCGGACTCGATCGGCGGGGTACACCATCGTGT 373
 Db 81 lGlyAlaAlaIleGlyIleTyrAsnAspLysPheGlyArgArgMetSerValLeu11 101
 QY 374 CGCGCGCTCATATCTTCGGGGGGSGTTCCTCATGGGTTCGCGCTCACTACGCCAT 433
 Db 101 eAlaAspValLeuPheLeuLeuGlyAlaLeuValMetValIleAlaHisalaProTrpVa 121
 QY 434 GCTCATGTCGGCGCTTCGTGGCGGCTCGCGGTGGGTACGGCTCATGTCGCGCC 493
 Db 121 lIleIleLeuGlyArgLeuLeuValGlyPheGlyValGlyMetAlaSerMetThrSerPr 141
 QY 494 GGTGTACACCGCGAGTGTCCGGCGTGGCGGTGGTCTTCACGCTCGTTCGCCGA 553
 Db 141 oLeuYrIleSerGluMetSerProAlaArgIleArgGlyAlaLeuValSerThrAsnGl 161

QY 554 GGTGTTTCATCAACTTCGGCATCTCTCGGTACCTCTCGAACTATGCTTCTCCGCTT 613
 Db 161 yLeuLeuIleThrGlyGlnPheLeuSerTyrLeuIleAsnLeuAlaPheValHisTh 181
 QY 614 GCGGTGAACCTCGGTGGCGCATCTCTCGGCATCGCGCGCGCGCTCGCTCGCTGCT 673
 Db 181 rProGlyThr-----TrpArgTrpMetLeuGlyValSerAlaIleProAlaIleleG1 199
 QY 674 CCGCTCATGTCGTGGCATCCCGAGTCGCGCGGTGGCTGCTCATCAAGAGGACGCT 733
 Db 199 nPheCysLeuMetLeuThrLeuProGluSerProArgTrpLeuTyrArgAsnAspArgly 219
 QY 734 CCGGACGCCAAGGTGGTCTGGAGAAG-----ACCTCCGACACGGCGGAGGAGCGCG 787
 Db 219 sAlaGluSerArgAspIleLeuGluArgIleTyrProAlaGluMetValGluAlaGlu11 239
 QY 788 GGAGCGCTTCGGCGCATCAAGCGCGCGCATCCCTGAGGAGCTCGACGCGACGT 847
 Db 239 eAlaAlaLeuLysGluSerValArgAlaGluThrAlaAspGluAspIleIleGlyHisTh 259
 QY 848 GGTGACCTCCCAAGAGAGGAGCGGAAACGAGAGCGGTGTGGAAGGAGCTCATCT 907
 Db 259 rPheSerAspLysLeuArgGlyAla-----Le 268
 QY 908 GTCCCGACCGCGCATCGCGCATCTCTGTCTCGGATCGGCATCCATCTCTTCCA 967
 Db 268 uSer---AsnProValValArgHisGlyLeuAlaGlyIleThrValGlnValAlaGl 287
 QY 968 GCATCGTGGCATCTCACTCGCTCTTCTACAGCCCTCTCGTGTCAAGAGCCCGG 1027
 Db 287 nGlnPheValGlyIleAsnThrValMetTyrTyrSerProThrIleLeuGlnPheAlaGl 307
 QY 1028 ATTAACGAAACGACCAACCTCTTGGCCACCATCTGGCGCTCGGTGTCTACCAAGAGCT 1087
 Db 307 yTyrAlaSerAsnLysThrAlaMetAlaLeuIleThrSerGlyLeuAsnAlaVa 327
 QY 1088 TTTCATCTGTGGGACTTTCTTCATCGACGGGTGTGGCGCGCGCTGTGTCTGGG 1147
 Db 327 lGlySerValValSerMetMetPheValAspArgTyrGlyArgLysLeuMetIle11 347
 QY 1148 CAGCAGCGGGGATAATCTCTCCCTCATCGGCTCGCGCGCGGCTCACCGCTCGTGG 1207
 Db 347 sSerMetPheGlyIleIleThrCysLeuValIleLeuAlaValPheAsnGluAlaSe 367
 QY 1208 CCAGCAC-----1214
 Db 367 rAsnHisalaProLysIleAspLysArgAspSerArgAsnPheAlaLysAsnAlaThrCy 387
 QY 1214 -----1214
 Db 387 sProAlaPheAlaProPheThrAlaSerArgSerProSerAsnTrpAsnCysMetLy 407
 QY 1215 -----CCGACGC 1222
 Db 407 sCysLeuGlnTyrAspCysGlyPheCysSerAsnGlyAlaGlnGluTyrAlaProGlyAl 427
 QY 1223 CAAGATA-----1229
 Db 427 aCysIleValGlnSerAlaAspMetLysAlaLeuCysHisSerLysGlyArgThrPhePh 447
 QY 1230 -----CCTTGGCGCATCGG-----CTAAGCATCGCTCCACCTCGCTACGT 1273
 Db 447 eLysAspGlyCysProSerLysPheGlyTyrLeuAlaIleValPheLeuGlyLeuTyrI1 467
 QY 1274 GCGCTTCTTCATCGCTTCGCGCTTCGCGCATCATCGTGGGTGTACAGCTCGGAGATCTTCCC 1333
 Db 467 eIleValTyrAlaProGlyMetGlyThrValProTrpIleValAsnSerGluIleTyrPr 487
 QY 1334 GCTCAGGTGGCGCTGGGCTGCTCGTGGGTGGCGCGCAACCGCTCACGAGCGG 1393
 Db 487 oLeuArgTyrArgGlyLeuAlaGlyIleAlaAlaValSerAsnTrpMetSerAsnLe 507


```
QY 1394 CGTCATCTCCATGACCTTCTGCTCGCTGCTCCAAAGCGCATCACCATCGCGCGACCTTCTT 1453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 507 uValValSerGluThrPheLeuThrThrAsnAlaValGlySerSerGlyThrPheLe 527
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1454 CCTTACTCCGGATCCGCGCGCTCGCTGGCTGTTCTTCTACACCTACTCCCGGAGAC 1513
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 527 uLeuPheAlaGlySerSerAlaValGlyLeuPhePheIleTrpLeuLeuValProGluTh 547
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1514 CCGCGGCGGACGCTGGAGGAGATGAGCAAGCTG 1547
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 547 rLysGlyLeuGlnPheGluGluValGluLysLeu 558
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 15
S47089
arabinose-proton symporter - Klebsiella oxytoca
C:Species: Klebsiella oxytoca
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C:Accession: S47089
R:Shatwell, K.P.; Charalambous, B.M.; McDonald, T.P.; Henderson, P.J.F.
submitted to the EMBL Data Library, June 1994
A:Description: The nucleotide sequence of the gene araE for arabinose-proton symport in
A:Reference number: S47089
A:Accession: S47089
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-472 <SRA>
A:Cross-references: EMBL:X79598; NID:g498919; PIDN:CAA56110.1; PID:g498920
C:Superfamily: Glucose transport protein
C:Keywords: arabinose transport

Alignment Scores:
Pred. No.: 1,31e-26 Length: 472
Score: 584.50 Matches: 143
Percent Similarity: 52.25% Conservative: 101
Best Local Similarity: 30.62% Mismatches: 182
Query Match: 16.17% Indels: 41
DB: 2 Gaps: 10

US-10-051-902A-21 (1-2017) x S47089 (1-472)
QY 168 GCCATCTCTCGCTCCATGACCTCCATCTCTCGCTACGATATCGGGGTGATGAGCGG 227
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 26 SerIleAlaAlaValAlaGlyLeuLeuPheGlyLeuAspIleGlyValIleAlaGly 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 228 GCGTGGTGTACATCAGAGGACTTCAACATAGTACGGGAAGTGGAGGTCTCTCATG 287
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 46 AlaLeuProPheIleThrAspHisPheValLeuSerSerArgLeuGlnGluTrpValVal 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 288 GGCATCTGAACCTCTACTCGCTCATCGCTCCCTCGCGGGCGGACGTCGGACTGG 347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 66 SerSerMetMetLeuGlyAlaAlaIleGlyAlaLeuPheAsnGlyTrpLeuSerPheArg 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 348 ATCGCGCGGCGGTACACCATCGTGTTCGCGCGCTCATATCTTCGCGGGGGTTCCTC 407
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 86 LeuGlyArgLysTyrSerLeuMetValGlyAlaValLeuPheValAlaGlySerValGly 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 408 ATGGGGTTCGCGCACTACGCTCATCGCTCATGTTCCGCGCTTCGCGCGCATCGGC 467
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 106 SerAlaPheAlaThrSerValGluuMetLeuValAlaArgIleValLeuGlyValAla 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 468 GTGGGCTACGGCTCATGATCGCGCGGTGTACACCGCGAGGTGTGCGCGGTCTCGCG 527
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 126 ValGlyIleAlaSerTyrThrAlaProLeuTyrLeuSerGluMetAlaSerGluAsnVal 145
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 528 CGTGGCTTCCTGAGTGTCTCCGAGGTGTTCATCACTTCGGCATCTCTCGGCTAC 587
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 146 ArgGlyLysMetIleSerMetTyrGlnLeuMetValThrLeuGlyIleValMetAlaPhe 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 588 GTCTCGAACTATGCTTCTCCCGCTGTCGCGGTGAACCTCGGTGGCGCATCATCTCGCG 647
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 166 LeuSerAspThrAlaPheSer-----TyrSerGlyAsnTrpArgAlaMetLeuGly 182
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 648 ATCGGCGGCGCGCTCGTCTGCTCGCTCATGCTGCTCGCATGCTCGGATCGCGG 707
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

Search completed: June 30, 2004, 18:53:07
Job time : 65 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:37:09 ; Search time 29.5 Seconds
(without alignments)
7120.373 Million cell updates/sec

Title: US-10-051-902A-21

Perfect score: 3614
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaaaa 2017

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp
-Q/cg2_1/US10051902/runat_30062004_164722_20400/app_query.fasta_1.2183
-DB=SwissProt_42 -QFMT=fastan -SUFFIX=resp -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10051902 @CGN 1.1 22 @runat_30062004_164722_20400 -NCPUs=6 -ICPU=3
-NO MVAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	654.5	18.1	629	1 MYCT HUMAN	Q96q22 homo sapien
2	609.5	16.9	461	1 CSBC BACSU	P46333 bacillus su
3	601	16.6	557	1 ITR2 SCHPO	P87110 schizosacch
4	584.5	16.2	472	1 ARAE KLEBX	P45598 klebsiella
5	582	16.1	457	1 XYL1 LACBR	O52733 lactobacill
6	579.5	16.0	472	1 ARAE ECOLI	P09830 escherichia
7	576	15.9	584	1 ITR1 YEAST	P30605 saccharomyc
8	574	15.9	491	1 XYLE ECOLI	P09098 escherichia
9	572.5	15.8	464	1 GALP ECOLI	P37021 escherichia
10	572.5	15.8	612	1 ITR2 YEAST	P30606 saccharomyc
11	555.5	15.4	482	1 YPIG BACSU	P54723 bacillus su
12	543.5	15.0	522	1 STP1 ARATH	P23586 arabidopsis
13	539	14.9	534	1 HUP3 CHLKE	Q29525 chlorella k
14	538.5	14.9	534	1 HUP1 CHLKE	P15686 chlorella k
15	531.5	14.7	468	1 GLCP SYNY3	P15729 synechocyst
16	525.5	14.5	464	1 ARAE BACSU	P96710 bacillus su
17	525	14.5	575	1 ITR1 SCHPO	Q10286 schizosacch
18	523.5	14.5	547	1 GTR1 LEIDO	Q01440 leishmania

19	505.5	14.0	510	1 HEX6 RICCO	Q07423 ricinus com
20	503.5	13.9	523	1 STC RICCO	Q41144 ricinus com
21	502	13.9	522	1 HTX RICCO	Q10710 ricinus com
22	486	13.4	592	1 HXT5 YEAST	P38695 saccharomyc
23	484	13.4	541	1 GT10 HUMAN	O95528 homo sapien
24	483.5	13.4	541	1 HXT2 YEAST	P23585 saccharomyc
25	482.5	13.4	490	1 GTR1 CHICK	P46896 gallus gall
26	475	13.1	477	1 GTR8 MOUSE	Q91163 mus musculus
27	468.5	13.0	570	1 HXT1 YEAST	P34465 saccharomyc
28	467.5	12.9	537	1 QAY NEUCR	P11636 neurospora
29	465.5	12.9	496	1 GTR3 CHICK	P28568 gallus gall
30	463	12.8	540	1 HUP2 CHLKE	Q39524 chlorella k
31	460.5	12.7	473	1 GLF ZYMO	P21906 zymomonas m
32	460.5	12.7	492	1 GTR1 RABIT	P13355 oryctolagus
33	456.5	12.6	492	1 GTR1 RAT	P11167 rattus norv
34	455.5	12.6	478	1 GTR8 RAT	Q9JJZ1 rattus norv
35	455.5	12.6	492	1 GTR1 MOUSE	P17809 mus musculus
36	455.5	12.6	569	1 HXT8 YEAST	P40886 saccharomyc
37	454.5	12.6	492	1 GTR1 HUMAN	P11166 homo sapien
38	454.5	12.6	567	1 RAG1 KLULA	P18631 kluyveromyc
39	453	12.5	477	1 GTR8 HUMAN	Q92V64 homo sapien
40	452.5	12.5	492	1 GTR1 BOVIN	P27674 bos taurus
41	451	12.5	546	1 HXT0 YEAST	P43581 saccharomyc
42	451	12.5	566	1 KHT2 KLULA	P53387 kluyveromyc
43	448	12.4	509	1 GTR4 HUMAN	P14672 homo sapien
44	445.5	12.3	495	1 GTR3 CANFA	P47842 canis famil
45	445.5	12.3	567	1 HXT9 YEAST	P40885 saccharomyc

ALIGNMENTS

RESULT 1	MYCT HUMAN	STANDARD;	PRT;	629 AA.
ID	Q96q22			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Proton myo-inositol co-transporter (Hmit).			
GN	SLC2A13			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21391710; PubMed=11500374;			
RA	Uldry M., Ibberson M., Horisberger J.-D., Chatton J.-Y.,			
RA	Riederer B.M., Thorens B.;			
RT	"Identification of a mammalian H(+)-myo-inositol symporter expressed predominantly in the brain."			
RL	EMBO J. 20:4467-4477(2001).			
CC	!- FUNCTION: H(+)-myo-inositol co-transporter. Can also transport related stereoisomers.			
CC	!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	!- TISSUE SPECIFICITY: Predominantly expressed in the brain.			
CC	!- PTM: Glycosylated.			
CC	!- SIMILARITY: Belongs to the sugar transporter family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AJ315644; CAC51116.1; --			
DR	Genew; HGNC:15956; SLC2A13.			
DR	InterPro; IPR007114; MFS.			
DR	InterPro; IPR005828; Sub transporter.			
DR	InterPro; IPR005829; Sug transporter.			
DR	InterPro; IPR003663; Sugar_transp.			

Qy 1248 ---AGCATCGCTCCACCCCTCGCTAGCTGCCTTCTTCCATCGCGCTTGGCCCATC 1304
 Db 489 ThrAlaLeuLeuGlyLeuLeuLeuLeuValPhePheAlaArgSerThrGlyProMet 508
 Qy 1305 ACGRGGGTGTACAGCTCGGAGATCTCCCGCTCCAGGTGGCGGTGGCC-----TGC 1358
 Db 509 ProTrpThrValAsnSerGluLeuLeuTyProLeuTrpAlaArgSerThrGlyAsnAlaCys 528
 Qy 1359 TCGTCGGGTGCGCGCCACCGCGTCCACCGCGGTCCACCGCGGTATCTCCATGACCTTCCTGTCG 1418
 Db 529 SerSerGlyLeu-----AsnTrpIlePheAsnValLeuValSerLeuThrPheLeuHis 546
 Qy 1419 CTGTCCAAAGCCATCACCATCGCGGCGAGTCTTCTCTCTACTCCGSCATCGCGGCGTC 1478
 Db 547 ThrAlaGluTyLeuThrTyTyTyGlyAlaPhePheLeuTyAlaGlyPheAlaAlaVal 566
 Qy 1479 GCTGTGGTGTCTTCTACACTCTCCCGAGACCGCGCGCGAGCTGGAGGAGATG 1538
 Db 567 GlyLeuLeuPheThrLeuTyGlyCysLeuProGluThrLysGlyLysLeuGluGluLe 586
 Qy 1539 AGCAAGCTGTTCGGCGAC-----ACGGCGCGCGCTCGGAATCAGACGAG 1583
 Db 587 GluSerLeuPheAspAsnArgLeuCysThrCysGlyThrSerAspSerAspGlu 604
 RESULT 2
 CSBC_BACSU STANDARD; PRF; 461 AA.
 AC P46333; Q32289;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE Probable metabolite transport protein cscb.
 GN CSBC OR SS92BR OR BSU39810.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RC STRAIN=168 / BGSCA1;
 RX MEDLINE=96093926; PubMed=7584049;
 RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
 RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
 RL genome between the gnt and iol operons.";
 RN DNA Res. 2:61-69(1995).
 RN [2]
 RP REVISIONS.
 RA Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
 RA Azavedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun K., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmeron P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescean E., Pujic P., Purrelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takenaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Tarnier P., Tognoni A.,
 RA Tasato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subillis.";
 RN Nature 390:249-256(1997).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=9303315; PubMed=10376822;
 RA Akbar S., Lee S.Y., Boylan S.A., Price C.W.;
 RT "Two genes from bacillus subtilis under the sole control of the
 RT general stress transcription factor sigmaB.";
 RL Microbiology 145:1069-1078(1999).
 CC -!- FUNCTION: Could serve either a nutritional or an osmotic
 CC protection function.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- SIMILARITY: Belongs to the sugar transporter family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB005554; BAA21604.1; -;
 DR EMBL; Z99124; CAB16017.1; -;
 DR PIR; D70073; D70073.
 DR Subtilist; BG11360; cscb.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR Pfam; PF00083; sugar tr; 1.
 DR PRINTS; PR00171; SUGTRANSPORT.
 DR TIGRfam; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Transmembrane; Complete proteome.
 FT DOMAIN 1 14
 FT TRANSMEM 15 35
 FT DOMAIN 36 38
 FT TRANSMEM 39 59
 FT DOMAIN 60 76
 FT TRANSMEM 77 97
 FT DOMAIN 98 104
 FT TRANSMEM 105 125
 FT DOMAIN 126 139
 FT TRANSMEM 140 160
 FT DOMAIN 161 163
 FT TRANSMEM 164 184
 FT DOMAIN 185 241
 FT TRANSMEM 242 262
 FT DOMAIN 263 280
 FT TRANSMEM 281 301
 FT DOMAIN 302 308
 FT TRANSMEM 309 329
 FT DOMAIN 330 341
 FT TRANSMEM 342 362
 FT DOMAIN 363 378
 FT TRANSMEM 379 399
 FT DOMAIN 400 402
 FT TRANSMEM 403 423
 FT DOMAIN 424 461
 FT TRANSMEM 461 400
 FT CONFLICT 461 AA; AD59CB5DD6372FE0 CRC64;
 SQ SEQUENCE

Alignment Scores:

Pred. No.: 2,5e-22 Length: 461
 Score: 609.50 Matches: 149
 Percent Similarity: 49.59% Conservative: 93
 Best Local Similarity: 30.53% Mismatches: 195
 Query Match: 16.86% Indels: 51
 DB: 1 Gaps: 11

US-10-051-902A-21 (1-2017) x CSBC_BACSU (1-461)

QY 180 TCCATGACCTCATCTCTCGGTAGATATCGGGGTGATGAGCGGGCGTCTGCTAC 239
 DB 15 AlaleuGlyLeuLeuTyGlyThrAspThrGlyValIleSerGlyAlaLeuLeuPhe 34
 QY 240 ATCAAGAGGACTTCAACATCAGTCAGGAGGAGTTCATCGGCATCTACTGAAC 299
 DB 35 IleAsnAspIleProLeuThrThrLeuThrGluGlyLeuValSerMetLeuLeu 54
 QY 300 CTCCTACTCGCTCATCGCTCTCGCGCGGGCGGAGCTCGACTGATCGCGCGCG 359
 DB 55 LeuGlyAlaIlePheGlySerAlaLeuSerGlyThrCysSerAspArgTrpGlyArg 74
 QY 360 TACACCATCGTTCGCGCGCGTATATCTTCGCGGGGGGTCTCTCATGGGGTTCGCC 419
 DB 75 LysValValPheValLeuSerIleIlePheIleIleGlyAlaLeuAlaCysAlaPheSer 94
 QY 420 GTCAACTAGCCATGCTCATGTTTCGCGCTTCGTCGCGGCATCGCGGTGGCTACGCG 479
 DB 95 GlnThrIleGlyMetLeuIleAlaSerArgValIleLeuGlyLeuAlaValGlyIle 114
 QY 480 CTCATGATCGCGCGTGTACACCGCGAGTGTGCGCGGTGCGCGGTGGCTTCCTG 539
 DB 115 ThrAlaLeuValProValGlyLeuSerGluMetAlaProThrLysIleArgGlyThrLeu 134
 QY 540 ACGTGTTCGCGGAGTGTTCATCAACTTCGCGATCTGCTCGGGTACGTCTCGAATAT 599
 DB 135 GlyThrMetAsnLeuMetIleValThrGlyIleLeuLeuAlaTyIleValAsnTyr 154
 QY 600 GCTTCTCCGCTGCTCCCTGAACCTCGGTGGCGCATCATGCTCGGCATCGCGCGCG 659
 DB 155 LeuPheThr-----ProPheGlu---AlaTrpArgTrpMetValGlyLeuAlaAlaVal 171
 QY 660 CGTCCGCTGCTCGCGCTCATGCTCGGATCGCGATCGCGAGTCCGCGCGGTGCTGTC 719
 DB 172 ProAlaValLeuLeuLeuIleGlyIleAlaPheMetProGluSerProArgTrpLeuVal 191
 QY 720 ATGAAGGACGCTCGCGAGCCAGTGTGTGTGAGAGAGACTCCGACACGCGCGAG 779
 DB 192 LysArgGlySerGluGluAlaArgIleMetAsnIleThrHisAsp----- 208
 QY 780 GAGGCGCGGAGCGCTCGCGCATCATCAAGCCCGCGCGCATCTCGAGGAGCTCGAC 839
 DB 209 -----ProLysAspIleGlu 213
 QY 840 GCGACGTGTGTACCGTCCCAAGAGAGGAGCGGAAACGAGAGCGGGTGTGAAGAG 899
 DB 214 MetGluLeuAlaGluMet---LysGlnGlyAlaGluLys-----LysGlu 228
 QY 900 CTCATCTGTCCCCACCCCGGCC-----ATGCGGCGCATCTGCTCGGGATCGGC 953
 DB 229 ThrThrLeuGlyValLeuLysTrpIleArgProMetLeuLeuIleGlyValGly 248
 QY 954 ATCCATCTTTCAGACGTGTGGCATTCACCTCGGTGCTTCTACAGCCCTCTCGT 1013
 DB 249 LeuAlaIlePheGlnGlnAlaValGlyIleAsnThrValIleTyTrAlaProThrIle 268
 QY 1014 TTCAGAGCCCGGTTTAAAGACGACAAACACTTCTGGCCACCTTGGCGCTTCGGT 1073
 DB 269 PheThrLysAlaGlyLeuGlyThrSerAlaSerAlaLeuGly----ThrMetGlyIleGly 287
 QY 1074 CTCACCAAGAGGCTTTTCATCTTGTGGCGACTTCTTCATCGACGCGGTGCGCGCG 1133

DB 288 IleLeuAsnValIleValMetCysIleThrAlaMetIleLeuIleAspArgValGlyArgIys 307
 QY 1134 CGCTGTTCTGCGGACGACGCGGATATPCTCTCCCTCATCGGCTCGGCGCGGG 1193
 DB 308 LysLeuLeuIleTrpGlySerValGlyIleThrLeuSerLeuAlaLeuSerGlyVal 327
 QY 1194 CTCACGCTGCTCGGCGACGACCCCGACGACGATACCTTGGGCCATCGGCTTAAGCATC 1253
 DB 328 LeuLeuThrLeuGlyLeu-----SerAlaSerThrAlaTrp-----MetThrVal 342
 QY 1254 GCCTCCACCTCGCTCATGCTCGCTTCTTCTCCATCGGCTTGGCCCATCACGTGGGTG 1313
 DB 343 ValPheLeuGlyValTyIleValPheTyTrGlnAlaThrTrpGlyProValValTrpVal 362
 QY 1314 TACAGTTCGAGATCTTCCGCTCCAGGTGCGCGCGCTGGGTGCTCGTGGCGCTCGCC 1373
 DB 363 LeuMetProGluLeuPheProSerLysAlaArgGlyAlaAlaThrGlyPheThrThrLeu 382
 QY 1374 GCCAACCGCTCACGACGCGCTCATCTCATGACCTTCTCTGCTGTCTCAAGGCCATC 1433
 DB 383 ValLeuSerAlaAlaAsnLeuIleValSerLeuValPheProLeuMetLeuSerAlaMet 402
 QY 1434 ACCATCGCGCGAGTCTTCTCTACTCCGCGCATCGCGCGCTCGCTGGGTGTCTTC 1493
 DB 403 GlyIleAlaTrpValPheMetValPheSerValIleCysLeuLeuSerPhePheAla 422
 QY 1494 TACACCTACTCCCGGAGACCGCGCGCGGAGGAG-----ATGAGC 1541
 DB 423 PheTyMetValProGluThrLysGlyLysSerLeuGluGluIleGluAlaSerLeuLys 442
 QY 1542 AAGCTGTTCGCGACACGCGCGCTCGGATCAGACGACGACGACGACGACGACGACG 1601
 DB 443 LysArgPhe-----LysLysLysLys 449
 QY 1602 AAGGTGGAATGCGCGCCACTAAC 1625
 DB 450 SerThrGlnAsnGlnValLeuAsn 457
 RESULT 3
 ID ITR2_SCHPO STANDARD; PRT; 557 AA.
 AC P87110; P78901;
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Myo-inositol transporter 2.
 GN ITR2 OR SPAC2098.03.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=968 h90;
 RX MEDLINE=9828265; PubMed=95560432;
 RA Niederberger C., Grab R., Schweingruber A.-M., Fankhauser H.,
 RA Rusu M., Poitelea M., Edenharter L., Schweingruber M.B.;
 RT "Exogenous inositol and genes responsible for inositol transport are
 RT required for mating and sporulation in Schizosaccharomyces pombe.";
 RL Curr. Genet. 33:255-261(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Henderson P.J.F.;
 RT "Mammalian and bacterial sugar transport proteins are homologous.";
 RL Nature 325:641-643(1987).
 [2]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / JM2433;
 RX MEDLINE=88228015; PubMed=2836407;
 RA Maiden M.C.J., Jones-Mortimer M.C., Henderson P.J.F.;
 RT "The cloning, DNA sequence, and overexpression of the gene araE
 coding for arabinose-proton symport in *Escherichia coli* K12.";
 RL J. Biol. Chem. 263:8003-8010(1988).
 [3]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
 Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 [4]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,
 Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 Grobeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,
 Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 [5]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 Kuhara S., Saiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL Nature Res. 8:11-22(2001).
 [6]
 RN
 RP PRELIMINARY SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=84114868; PubMed=6319708;
 RA Stoner C., Schleif R.F.;
 RT "The araE low affinity L-arabinose transport promoter. Cloning,
 sequence, transcription start site and DNA binding sites of
 regulatory proteins.";
 RL J. Mol. Biol. 171:369-381(1983).
 CC
 CC -!- FUNCTION: Uptake of arabinose across the boundary membrane with
 the concomitant export of a proton (symport system).
 CC
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC
 CC -!- SIMILARITY: Belongs to the sugar transporter family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sb-sib.ch).
 CC
 CC -----
 CC EMBL; J03732; AAA23469.1; -;
 CC EMBL; X00272; CAA25075.1; -;
 CC EMBL; U28581; ABA40488.1; -;
 CC EMBL; A0000368; AAC75880.1; -;
 CC EMBL; A0005513; AAG57953.1; -;
 CC EMBL; AP002563; BAB37121.1; -;
 CC PIR; B26430; B26430.
 CC PIR; B91091; B91091.

DR PIR; E85936; E85936.
 DR EcoGene; EGI0056; araE.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003683; Sugar_transpt.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PRO0171; SUGTRNSPORT.
 DR TIGRFAMS; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
 KW Complete proteome.
 FT DOMAIN 1 29 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 30 50 1 (POTENTIAL).
 FT DOMAIN 51 63 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 64 84 2 (POTENTIAL).
 FT DOMAIN 85 91 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 92 112 3 (POTENTIAL).
 FT DOMAIN 113 114 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 115 135 4 (POTENTIAL).
 FT DOMAIN 136 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 155 175 5 (POTENTIAL).
 FT DOMAIN 176 178 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 179 199 6 (POTENTIAL).
 FT DOMAIN 200 257 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 258 278 7 (POTENTIAL).
 FT DOMAIN 279 297 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 298 318 8 (POTENTIAL).
 FT DOMAIN 319 325 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 326 346 9 (POTENTIAL).
 FT DOMAIN 347 361 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 362 382 10 (POTENTIAL).
 FT DOMAIN 383 404 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 405 425 11 (POTENTIAL).
 FT DOMAIN 426 427 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 428 446 12 (POTENTIAL).
 FT DOMAIN 447 472 CYTOPLASMIC (POTENTIAL).
 FT CONFLICT 26 28 SVA -> YDR (IN REF. 6).
 SQ SEQUENCE 472 AA; 51684 MW; 411990A41D44393 CRC64;
 Alignment Scores:
 Pred. No.: 6.77e-21 Length: 472
 Score: 579.50 Matches: 146
 Percent Similarity: 51.23% Conservative: 105
 Best Local Similarity: 29.88% Mismatches: 192
 Query Match: 16.03% Indels: 47
 DB: 1 Gaps: 11
 US-10-051-902A-21 (1-2017) x ARAE_ECOLI (1-472)
 QY 102 TCQCGCGCGCTCGCGAGCGCTCGCGCGAAGAGAGAGGCGGTCGGTTCGCCTTC 161
 Db 8 SerAlaLeuThrProArgSerLeuArgAspThrArgArgMetAsnMetPheValSerVal 27
 QY 162 GCCTCGCGCATCTCGCGCTCCATCGACCTCCCTCCGCTACGATCGATATCGGGGTGATG 221
 Db 28 AlaAlaAlaVal-----AlaGlyLeuLeuPheGlyLeuAspIleGlyValIle 43
 QY 222 ACQGGGGCGCTGCTGATCATCAAGAGGAGCTTCAACATCAGTCAGCGGAGGTGGAGGTT 281
 Db 44 AlaGlyAlaLeuProPheIleThrAspHisPheValLeuThrSerArgLeuGlnIleTrp 63
 QY 282 CTCATGGGCATACCTCAACTCTCTACTCGCTCATCGGCTCTTCGCGGGCGGCGGACGTCG 341
 Db 64 ValValSerSerMetMetLeuGlyAlaAlaIleGlyAlaLeuPheAsnGlyTrpLeuSer 83
 QY 342 GACTGATCGCGCGCGGTACACCATCGTGTTCGCCGCCGTCATATTTCTCGCGGGGSG 401
 Db 84 PheArgLeuGlyArgLysTyrSerLeuMetAlaGlyAlaIleLeuPheValLeuGlySer 103
 QY 402 TTCCTCATGGGTTTCGCGGTCAACTACGCCATGCTCATGTCGCGCGCTTCGTCGCCGCG 461

Db 104 IleGlySerAlaPheAlaThrSerValGluMetLeuIleAlaAlaArgValValLeuGly 123
QY 462 ATCCGCGTGGCTACCGGCTCATGATCGCGGTGTACACCGCGAGGTGTGCGCGGCG 521
Db 124 IleAlaValGlyIleAlaSerTyThrAlaProLeuTyThrLeuSerGluMetAlaSerGlu 143
QY 522 TCGGCGGTGCTTCCTCCGCTGCTCCGAGGTGTTCATCACTTCGGCATCTCTC 581
Db 144 AsnValArgGlyLysMetIleSerMetTyThrGlnLeuMetValThrLeuGlyIleValLeu 163
QY 582 GGGTACGCTCGCACTATGCTTCTCCGCTGCGGTGCACTTCGGGTGCGCATCATG 641
Db 164 AlaPheLeuSerAspThrAlaPheSer-----TyrSerGlyAsnTyrArgAlaMet 180
QY 642 CTCGGCATCGCGCGGCGCTCGTGTCTCGCTCATGTCGTCGTCGTCGTCGTCGTCG 701
Db 181 LeuGlyValLeuAlaLeuProAlaValLeuLeuIleLeuValValPheLeuProAsn 200
QY 702 TCGCGCGGTGCTGCTCATGAAGGCGCTCGCGGACGCAAGTGTGCTCGAGAAG 761
Db 201 SerProArgTyrPheAlaGlyLysGlyArgHisIleGluAlaGluValLeuArgMet 220
QY 762 ACCTCCGACACGCGGAGCGCGGCGGCGCTGCGCGCATCAAGCGCGCGCGCGCG 821
Db 221 LeuArgAspThrSerGluLysAlaArgGluGluLeuAenGluIleArgGluSerLeuLys 240
QY 822 ATCCCTGAGGAGCTCGACGCGGCGCTGCTGTCGCTCCCGAGAGAGCGGGAACGAG 881
Db 241 Leu-----LysGlnGlyGly----- 245
QY 882 AAGCGGTGTGAAGAGCTCATCTGCTCCCGACCGCGGCGCATCGCGGCGCATCTCTG 941
Db 246 -----Trp-----AlaLeuPheLysIleAsnArgAsnValArgAlaValPhe 260
QY 942 TCCGGGATCGGATCCACTTCTTCAGATCGGTGGGATCTACTCGTCTCTCTAC 1001
Db 261 LeuGlyMetLeuLeuGlnAlaMetGlnGlnPheThrGlyMetAsnIleMetTyThr 280
QY 1002 ACCCTCTCTGTTCAAGACCGCGGATTAACGACGACAAACACTTCTGGGCGACCAT 1061
Db 281 AlaProArgIlePheLysMetAlaGlyPheThrThrThrGluGlnGlnMetIleAlaThr 300
QY 1062 TGGCGGTGCGGTGTCACCAAGAGGCTTTCATCTGTGGGACTTCTTCATCGACGCG 1121
Db 301 LeuValValGlyLeuThrPheMetPheAlaThrPheIleAlaValPheThrValAspLys 320
QY 1122 GTCCGCGCGCGCTGTTGTCGCGCAGCGCGGCGGATACTCTCTCCCTCATCGGC 1181
Db 321 AlaGlyArgLysProAlaLeu-----LysIleGlyPheSerValMetAlaLeu----- 336
QY 1182 CTCGCGCGCGGCTCACCGCTGCTGCGCAG-----CACCCC 1217
Db 337 -----GlyThrLeuValLeuGlyTyrcysLeuMetGlnPheAsnGlyThrAla 353
QY 1218 GACGCAAGATACCTTGGCCATCGGCTAGCATCGCTCCACCTCGCTAGTGGCGC 1277
Db 354 SerSerGlyLeuSerTrp-----LeuSerValGlyMetThrMetCysIleAla 370
QY 1278 TTCTCTCATCGGCTTGGCCCATCATCGTGGGTGTACAGTCGGAGATCTCCCGCTC 1337
Db 371 GlyTyAlaMetSerAlaAlaProValValTrpIleLeuCysSerGluIleGlnProLeu 390
QY 1338 CAGGTGCGGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1397
Db 391 LysCysArgAspPheGlyIleThrCysSerThrThrThrThrThrThrThrThrThrThr 410
QY 1398 ATCTCATCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1457
Db 411 IleGlyAlaThrPheLeuThrLeuLeuAspSerIleGlyAlaAlaGlyThrPheTrpLeu 430
QY 1458 TACTCGGCGATC---GCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1514

431 TyrThrAlaLeuAsnIleAlaPheValGlyIleThrPheTrp---LeuIleProGluThr 449
1515 CCGCGCGCGACCTGAGGAGATGAGCAAG 1544
450 LysAsnValThrLeuGluHisIleGluArg 459
RESULT 7
ID ITRI YEAST STANDARD; PRT; 584 AA.
AC P30605; 1993 (Rel. 25, Created)
DT 01-APR-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE MYO-INOSITOL TRANSPORTER 1.
GN ITRI OR YDR497C OR D9719.3.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91250431; PubMed=2040626;
RA Nikawa J.-I., Tsukagoshi Y., Yamashita S.;
RT "Isolation and Characterization of two distinct myo-inositol
transporter Genes of Saccharomyces cerevisiae.",
RL J. Biol. Chem. 266:11184-11191 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
Winant A., Yelton M., Botstein D., Davis R.W.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR TRANSPORTER FOR MYO-INOSITOL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the sugar transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; D90352; BAA14366.1; -.
CC EMBL; U33057; AAB64939.1; -.
CC PIR; S69555; S69555.
CC GERMOnline; 140989; -.
CC SGD; S0002905; ITRI.
CC GO; GO:0005365; Fmyo-inositol transporter activity; IMP.
CC GO; GO:0015798; Pmyo-inositol transport; IMP.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub transporter.
CC InterPro; IPR005829; Sug transporter.
CC InterPro; IPR003663; Sugar transp.
CC Pfam; PF00083; sugar tr_1.
CC PRINTS; PR00171; SUGTRNSPORT.
CC TIGRFAMS; TIGR00879; SP; 1.
CC PROSITE; PS00850; MFS; 1.
CC PROSITE; PS00216; SUGAR TRANSPORT 1; 2.
CC PROSITE; PS00217; SUGAR TRANSPORT 2; 1.
CC Transmembrane; Sugar transport; Glycoprotein.
CC DOMAIN 1 81 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 82 102 1 (POTENTIAL).
CC TRANSMEM 103 129 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 130 150 2 (POTENTIAL).
CC TRANSMEM 151 163 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 164 184 3 (POTENTIAL).
CC TRANSMEM 185 186 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 187 207 4 (POTENTIAL).

```

FT DOVAIN 208 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 5 (POTENTIAL).
FT DOVAIN 237 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 267 6 (POTENTIAL).
FT DOVAIN 268 349 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 350 370 7 (POTENTIAL).
FT DOVAIN 371 376 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 377 397 8 (POTENTIAL).
FT DOVAIN 398 400 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 401 421 9 (POTENTIAL).
FT DOVAIN 422 441 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 442 462 10 (POTENTIAL).
FT DOVAIN 463 486 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 487 507 11 (POTENTIAL).
FT DOVAIN 508 510 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 511 531 12 (POTENTIAL).
FT DOVAIN 532 584 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLNAC. 1) (POTENTIAL).
FT CONFLICT 43 44 TL -> HI (IN REF. 1).
SQ SEQUENCE 584 AA; 63569 MW; 42543E30A102DC65 CRC64;

Alignment Scores:
Pred. No.: 1,01e-20 Length: 584
Score: 576.00 Matches: 148
Percent Similarity: 50.40% Conservative: 106
Best Local Similarity: 29.37% Mismatches: 206
Query Match: 15.94% Indels: 44
DB: 1 Gaps: 11

US-10-051-902a-21 (1-2017) x ITR1_YEAST (1-584)
QY 153 TTGCGCTTCGCGCCATCCCTCGCTCCATGACCTCCATCCTCCTCGGCTACGATATC 212
DB 85 PheileithrLeuThrPheValAlaSerIleSerGlyPheMetPheGlyTyAspThr 104
QY 213 GGGGTGATGACGGGGCTCGCTGATCATCAAGAGAGACTTCAAC-----ATCAGT 263
DB 105 GlyTyIleSerAlaLeuIleSerIleGlyThrAspLeuAspHisIlyValLeuThr 124
QY 264 GACGGGAAGTGGAGGTCTCTATGGGCATACGAACTCTACTCGCTCATCGCTCCTTC 323
DB 125 TyGlyGlyGlyGluIleValThrAlaAlaThrSerLeuGlyAlaLeuIleThrSerIle 144
QY 324 GCGCGGGGGAGCTCGGACTGATCGCGGGGTACACCATCGCTGTGTGCGCGCGTC 383
DB 145 PheAlaGlyThrAlaAlaAspIlePheGlyArgIlySerGlyLeuMetGlySerAsnLeu 164
QY 384 ATATTCTCGCGGGGGTTCCTCATGGGTTCGCGTCACTACGCGCATGCTCATGTC 443
DB 165 MetPheValIleGlyAlaIleLeuGlnValSerAlaHisThrPheTrpGlnMetAlaVal 184
QY 444 GCGCGTTCGTGGCGGCATCGCGGTGCTAGCGCTCATGATCGCGCGGTGTACACC 503
DB 185 GlyArgLeuIleMetGlyPheGlyValGlyIleGlySerLeuIleAlaProLeuPheIle 204
QY 504 GCGAGTGTGCGCGGCTCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 563
DB 205 SerGluLeuAlaProIlyMetIleArgGlyArgLeuThrValIleAsnSerLeuTrpLeu 224
QY 564 AACTTCGCGCATCTCGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 611
DB 225 ThrGlyGlyGlnLeuValAlaTyGlyCysGlyAlaGlyLeuAsnTy----- 240
QY 612 TTGCGCGTGAACTCGGGTGGCGCATCATCGCGATCGCGCGCGCGCGCTCGGCTG 671
DB 241 -----ValAsnAsnGlyTyPheGlyLeuValGlyLeuSerLeuIleProThrAlaVal 258
QY 672 CTCGCGTCACTGCTCGCGCATCGCGAGTTCGCGGTGCTGCTGCTGCTGCTGCTGCTG 731
DB 259 GlnPheThrCysLeuCysPheLeuProAspThrProArgTyTyValMetIlyGlyAsp 278
QY 732 CTCGCGGACGCCAAGTGTGCTGCGGAGAACG---TCCGACACGCGGAGGAGCGCGG 788

```

```

Db 279 LeuAlaArgAlaThrGluValLeuIlySerTyThrAspThrSerGluGluIleIle 298
QY 789 GAGCGCTTGGCGGACATCAAGCCCGCGCATCCCTGAGAGCTCGACGGCGACGTG 848
Db 299 GluArgLysValGlu-----GluLeu 305
QY 849 GTGACCGTCCCCAAGAGA-----GGAGCGAAACGAGACGGGTGTGGAGAGCTC 902
Db 306 ValThrLeuAsnGlnSerIleProGlyLysAsnValProGluLysValTrpAsnThrIle 325
QY 903 ---ATCCTCTCCCGACCCCGCCATCGCGCATCTCTCTCGGGATCGGCATCCAC 959
Db 326 LysGluLeuHisThrValProSerAsnLeuArgAlaLeuIleIleGlyCysGlyLeuGln 345
QY 960 TTCCTTCAGACGCTTGGGCAATCACTCGCTGCTCTACAGCCCTCTCGTGTCAAG 1019
Db 346 AlaIleGlnGlnPheThrGlyTrpAsnSerLeuMetTyPheSerGlyThrIlePheGlu 365
QY 1020 AGCCCGGATTAACGAACGACAAACACTTCTTGGGCACACACTTGGCGCTCGGTGTCAC 1079
Db 366 ThrValGlyPheLysAsnSer-----AlaValSerIleIleValSerGlyThr 382
QY 1080 AAGAGCTTTTCATCTTGTGGCACTTCTTCATCGACGGGTGCGGGCGCGCGCTG 1139
Db 383 AsnPheIlePheThrLeuValAlaPhePheSerIleAspLysIleGlyArgThrIle 402
QY 1140 TTGCTGGGCGACGACGGCGGGATATCTCTCCCTCATCGCTCGCGCTCGCGCGCTCACC 1199
Db 403 LeuLeuIleGlyLeuProGlyMetThrMetAlaLeuValCysSerIleAlaPheHis 422
QY 1200 GTCGCGCGCACGACCCCGCGCGCAAGATA-----CCT 1232
Db 423 PheLeuGlyIleLysPheAspGlyAlaValAlaValValSerSerGlyPheSerSer 442
QY 1233 TGGCGCATCGGCTAAGCATCGCTCCACCCCTAGCTCGCTTCTTCTTCCATCGCG 1292
Db 443 TrpGlyIle---ValIleValPheIleValPheAlaAlaPheTyAlaLeuGly 461
QY 1293 CTGCGCCCATCACGCTGGGTGTACAGCTCGGAGATCTTCCGCTCCAGTTCGCGCGCTG 1352
Db 462 IleGlyThrValProTrp---GlnGlnSerGluLeuPheProGlnAsnValArgGlyIle 480
QY 1353 GGCTCTCGCTCGGCTCGCGCGCAACCGCTGACAGCGGCTCATCTCCATGACCTTC 1412
Db 481 GlyThrSerTyAlaThrAlaThrAsnTrpAlaGlySerLeuValIleAlaSerThrPhe 500
QY 1413 CTGCTCGCTCAAGGCCATCAGCATCGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1472
Db 501 LeuThrMetLeuGlnAsnIleThrProAlaGlyThrPheAlaPheAlaGlyLeuSer 520
QY 1473 GCGCTCGCTCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1532
Db 521 CysLeuSerThrIlePheCysTyPheCysTyPheProGluLeuSerGlyLeuGluLeu 540
QY 1533 GAGATGACCAAGCTTTCGCGGACACGGCCCGCTCGGAATCAGACGACCGACCAAG 1592
Db 541 GluValGlnThrIleLeuLysAspGlyPheAsnIleLysAlaSerLysAlaLeuAlaLys 560
QY 1593 GAGAAGAAGAG 1604
Db 561 LysArgLysGln 564

RESULT 8
XLYE ECOLI
ID XLYE ECOLI STANDARD; PRT; 491 AA.
AC P09058;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D-xylose-proton symporter (D-xylose transporter).
GN XLYE OR B4031 OR Z5629 OR ECS5014.
OS Escherichia coli, and
OS Escherichia coli O157:H7.

```



```
QY 255 ---AACATCAGTACGGGAAGGTGGAGTTCTCATGGGCATACCTGAACTCTACTCGCTC 311
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 GlnAsnLeuSerGluSerAlaAlaAsnSerLeuLeuGlyPheCysValAlaSerAlaLeu 65
QY 312 ATCGGCTCTTCGGCGGGCGGGCGG-----ACGTCCGAGTCATCGGCGGGCGG 359
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 IleGlyCysIleIleGlyAlaLeuGlyGlyTyrCysSerAsnArgPheGlyArgArg 85
QY 350 TACACCATCGTTCGCGCGCGCTCATATCTTCCTCCGGGG----- 398
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 AspSerLeuIleAlaAlaValLeuPhePheIleSerGlyValGlySerAlaTyrPro 105
QY 399 ----- 416
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 GluLeuGlyPheThrSerIleAsnProAspAsnThrValProValTyrLeuAlaGlyTyr 125
QY 417 GCCTGCAACTACGCATCTCATGTCGGCGCTTCGTGGCGGCATCGCGTGGGTAC 476
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 ValProGluPheValIleTyr-----ArgIleIleGlyGlyIleGlyValGlyLeu 142
QY 477 GCGTCTATGATCGCGCGGTGTACACCGCGAGGTGTGCGCGGTGCGCGGTGCTTC 536
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
143 AlaSerMetLeuSerProMetTyrIleAlaGluLeuAlaProAlaHisIleArgGlyLys 162
QY 537 CTGACGTCGTTCGCGGAGGTTCATCAACTTCGCGCATCTGCTCGGCTACGTCGAA 596
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 LeuValSerPheAsnGlnPheAlaIleIlePheGlyGlnLeuLeuValTyrCysValAsn 182
QY 597 TATGCTTCTCCCGCTTCGCG-----CTGAACCTC-----GGGTGGCGCATCATG 641
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
183 TyrPheIleAlaArgSerGlyAspAlaSerTrpLeuAsnThrAspGlyTyrArgTyrMet 202
QY 642 CTCGGCATCGCGCGCGCTCGCTCGCTCGCTCATGTCGCTCGCATGCTCGCGCGGAG 701
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 PheAlaSerGluCysIleProAlaLeuLeuPheLeuMetLeuLeuTyrThrValProGlu 222
QY 702 TCGCGCGGTGCTGTCATGAAGGACGCTCGCGACGCCAAGTGTGCTCGGAGAG 761
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
223 SerProArgTrpLeuMetSerArgGlyLysGlnGluAlaGluGlyIleLeuArgLys 242
QY 762 ACC-----TCGACACGCGGAGGAGCGCGGCGCTGCGCCGACATCAAGCGCGCC 815
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
243 IleMetGlyAsnThrLeuAlaThrGlnAlaValGlnGlu----- 255
QY 816 GCGGCGATCTCGTAGGAGCTCGACGGCGACGTGTGACCGTCCCGACGAGGAGCGGA 875
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 -----IleLysHisSerLeuAspHisGly-----ArgLysThrGly 267
QY 876 AACGAGAGCGGTGTGGAAGAGCTCATCTGTCCCGACCGCGCCATCGCGCGCATC 935
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 Gly-----ArgLeu 270
QY 936 CTGCTGTCGGGATCGGCATC-----CACCTTCTTCAGCATCGG 974
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
271 LeuMetPheGlyValGlyValIleValIleGlyValMetLeuSerIlePheGlnPhe 290
QY 975 TTGGGATTCATCTCGCTCTTCTACAGCCCTCTCGTGTCAAGACCGCCGATTAACG 1034
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
291 ValGlyIleAsnValValLeuTyrTyrAlaProGluValPheLysThrLeuGlyAlaSer 310
QY 1035 AACGACAAACACTTCTTGGGCACACTTGGCGCTCGCTGCTCACCAGAGCTTTTCATC 1094
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
311 ThrAspIleAlaLeuLeuGlnThrIle---IleValGlyValIleAsnLeuThrPheThr 329
QY 1095 TTGTGGCGATTTCTTCTCATCGACGCGGTGCGGCGCGCTGTGTGCGGCGAGCAGC 1154
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
330 ValLeuAlaIleMetThrValAspLysPheGlyArgLysProLeuGln----- 345
QY 1155 GCGGGAATATCTCTCCCTCATCGGCTCGCGCGCGGTGTACCGTGTGCGCGAGCAC 1214
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 -----IleIleGlyAlaLeuGlyMetAlaIleGlyMetPheSerLeuGlyThrAla 362
QY 1215 CCGGACGCCAAGATACCTTGGCGCATCGCGCTAAGCATCGCTCCACCTCGCTACGTC 1274
```

```
Db 363 PheTyrThrGlnAlaProGlyIleValAlaLeu-----LeuSerMetLeuPheTyrVal 380
QY 1275 GCCTTCTTCTCATCGGCTTCGCGCCATCATCGTGGGTGTACAGTCGCGAGATCTTCCCG 1334
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
381 AlaAlaPheAlaMetSerTrpGlyProValCysTrpValLeuLeuSerGluIlePhePro 400
QY 1335 CTCACAGTGGCGCGCTGGCTGCTGCTGCGCGTGGCGCGCCAAACCGCTACCGCGCGC 1394
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
401 AsnAlaIleArgGlyLysAlaLeuAlaIleAlaValAlaAlaGlnTrpLeuAlaAsnTyr 420
QY 1395 GTCATCTCCATGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 PheValSerTrpThrPheProMetMetAspLysAsnSerTrpLeuValAlaHisPheHis 440
QY 1440 ---GGCGGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1496
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
441 AsnGlyPheSerTyrTrpIleTyrGlyCysMetGlyValLeuAlaAlaLeuPheMetTrp 460
QY 1497 ACCTACCTCCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1556
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
461 LysPheValProGluThrLysGlyLysThrLeuGluLeuGluAlaLeuTrp----- 478
QY 1557 ACGGCGCGCGCTCGGATCATCAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1616
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
479 -----GluPro-----GluThrLysLysThrGlnThr 488
QY 1617 GCACCT 1622
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
489 AlaThr 490

RESULT 9
GALP_ECOLI STANDARD; PRT; 464 AA.
ID GALP_ECOLI
AC P37021;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Galactose-proton symporter (Galactose transporter).
GN GALP OR B2943 OR C3529.
OS Escherichia coli, and
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992;
RN [1]_TaxID=562, 217992;
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA Roberts P.E.;
RL Thesis (1992), University of Cambridge, U.K.
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426517; PubMed=9278503;
RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH
THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
```

CC -!- SIMILARITY: Belongs to the sugar transporter family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U28377; AAA69110.1; --
 DR EMBL; AAC00377; AAC75980.1; --
 DR EMBL; AE016766; AAN81977.1; ALT_INIT.
 DR PIR; F65079; F65079.
 DR EcoGene; EG12148; galP.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR005829; Sug transporter.
 DR InterPro; IPR003663; Sugar transpt.
 DR Pfam; PF00883; sugar_trf_1.
 DR PRINTS; PRO0171; SUGKTRNSPRT.
 DR TIGRFAMs; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
 KW Complete proteome.
 FT DOMAIN 1 15
 FT TRANSMEM 16 36
 FT DOMAIN 37 56
 FT TRANSMEM 57 77
 FT DOMAIN 78 84
 FT TRANSMEM 85 105
 FT DOMAIN 106 112
 FT TRANSMEM 113 133
 FT DOMAIN 134 139
 FT TRANSMEM 140 160
 FT DOMAIN 161 171
 FT TRANSMEM 172 192
 FT DOMAIN 193 250
 FT TRANSMEM 251 271
 FT DOMAIN 272 290
 FT TRANSMEM 291 311
 FT DOMAIN 312 321
 FT TRANSMEM 322 342
 FT DOMAIN 343 351
 FT TRANSMEM 352 372
 FT DOMAIN 373 394
 FT TRANSMEM 395 415
 FT DOMAIN 416 416
 FT TRANSMEM 417 437
 FT DOMAIN 438 464
 SQ SEQUENCE 464 AA; 50982 MW; 07B08935BD8E3F8E CRC64;

Alignment Scores:

Pred. No.: 1,46e-20 Length: 464
 Score: 572.50 Matches: 141
 Percent Similarity: 49.79% Conservative: 99
 Best Local Similarity: 29.25% Mismatches: 205
 Query Match: 15.84% Indels: 37
 DB: 1 Gaps: 8

US-10-051-902A-21 (1-2017) x GALP_ECOLI (1-464)
 QY 111 CTGCGGAGCGCTCGCGCCGAAGAAGAGGCG-----AACGTCCGGTTCCGCTTCGCC 164
 Db 1 MetProAspAla-----LysLysGlnGlyArgSerAsnLysAlaMetThrPhePhe 17
 QY 165 TGGGCATCTCGCTCCATGACCTCCATCCCTCGGTACGATATCGGGGTGATGAGC 224
 Db 18 ValCysPheLeuAlaLeuAlaGlyLeuPheGlyLeuAspLysValleAla 37
 QY 225 GGGGGCTCGCTGTATCATCAAGAAGGACTTCACATCATGTCAGCGGAAGTGGAGTTCTC 284

Db 38 GlyAlaLeuProPheAlaLeuAspGluPheGlnIleThrSerHisThrGlnGluTrpVal 57
 QY 285 ATGGGCATACGAACCTCTACTCGCTCATCGGCTCTTTCGGCGGGGGGACGTCGAC 344
 Db 58 ValSerSerMetMetPheGlyAlaAlaValGlyAlaValGlySerGlyTrpLeuSerPhe 77
 QY 345 TGGATCGGCGCGGTACACCATCGTTCGCGCGGTGCATATCTTCGCGGGGSGTTC 404
 Db 78 LysLeuGlyArgLysSerLeuMetIleGlyAlaIleLeuPheValAlaGlySerLeu 97
 QY 405 CTATCGGGTTCGCGTCAACTACCCATGCTCATGTTGGCGCGCTTCGTGGCGGCATC 464
 Db 98 PheSerAlaAlaAlaProAsnValGluValLeuIleLeuSerArgValLeuGlyLeu 117
 QY 465 GCGGTGGCTACGCGCTCATGATCGCGCGGTGTACACCGCGAGGTGTCCGCGCGTCG 524
 Db 118 AlaValGlyValAlaSerTyrThrAlaProLeuTyrLeuSerGluIleAlaProGluLys 137
 QY 525 GCGGTGGCTTCCTGACGTTCGCGCGGTTCATCAACTTCGCGCATCTCTGCTCGGG 584
 Db 138 IleArgGlySerMetIleSerMetTyrGlnLeuMetIleThrIleGlyIleLeuGlyAla 157
 QY 585 TAGCTCTGAATATGCTTTCTCCGCTTCGCGCTGAACCTCGGCTGGCGCATCATGTC 644
 Db 158 TyrLeuSerAspThrAlaPheSer-----TyrThrGlyAlaTrpArgTrpMetLeu 174
 QY 645 GGCATCGGCGCGCGCTCGCTCGGCTCATGTCGCGCTCATGTCGCGCATCGCGAGTCG 704
 Db 175 GlyValIleIleIleProAlaIleLeuLeuLeuIleGlyValPhePheLeuProAspSer 194
 QY 705 CCGCGTGGCTGATGAAGGAGCGCTCGCGCGGTGTACAGGTGTCTGGAGAAGACC 764
 Db 195 ProArgTrpPheAlaAlaLysArgPheValAspAlaGluArgValLeuLeuArgLeu 214
 QY 765 TCCGACACGGGAGGAGCGCGCGGTGTGCGCGCATCAAGCGCGCGCGCGCATC 824
 Db 215 ArgAspThrSerAlaGluAlaLysArgGluLeuAspGluIleArgGluSerLeuGlnVal 234
 QY 825 CTGAGGAGCTCGACGGGACGTGTGACCTGCCAAGAGAGGAGGAGGAAACAGAGAG 884
 Db 235 LysGln-----SerGly----- 238
 QY 885 CCGGTGTGAAGAGCTCATCTGTCGCCGCGCATCGCGCGCATCTCTGCTGCC 944
 Db 239 -----Trp-----AlaLeuPheLysGluAsnSerAsnPheArgArgAlaValPheLeu 254
 QY 945 GGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCTACTCGCTCTTCTACAGC 1004
 Db 255 GlyValLeuLeuGlnValMetGlnGlnPheThrGlyMetAsnValIleMetTyrTyrAla 274
 QY 1005 CTTCTCGTTTCAAGAGCCCCGGATTAAAGACGCAACACTTCTTGGGCACCTTGG 1064
 Db 275 ProLysIlePheGluLeuAlaGlyTyrThrAsnThrThrGluGlnMetTrpGlyThrVal 294
 QY 1065 CCGTTCGCTGCACCAAGAGGCTTTCATCTGTTGGCGACTTTCCTTCATCGACGCGCTC 1124
 Db 295 IleValGlyLeuThrAsnValLeuAlaThrPheIleAlaIleGlyLeuValAspArgTrp 314
 QY 1125 GGGCGGCGCGCTGTGTGTCGGGACGACGCGGGGGGATATCTCTCCCTCATCGGCCTC 1184
 Db 315 GlyArgLysProThrLeu-----ThrLeuGlyPheLeuValMet 327
 QY 1185 GGGCGGCGCTCACCGTCTCGCGC-----CAGCACCCCGCGCCCAAGATACCTTGGGCG 1238
 Db 328 AlaAlaGlyMetGlyValLeuGlyThrMetMetHisIleGlyIleHisSerProSerAla 347
 QY 1239 ATCGGCTTAAGCATCGCTCCACCTCGCTCGCTCGCTCTTCTTCCATCGGCCTTGGC 1298
 Db 348 GlnTyr-PheAlaIleAlaMetLeuLeuMetPheIleValGlyPheAlaMetSerAlaGly 367
 QY 1299 CCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGTGTGCGCGCTGGCTGC 1358

Db 368 ProLeu1etpValLeuCySerGlu1edlnProLeuLysGlyArgAspPheGlyIle 387
Qy 1359 TCGTCGGCGTCGCGCCACCGGTCACCGAGCGGTCATCTCCATGACCTTCCTGTCG 1418
Db 388 ThrCysSerThrAlaThrAsnTrpIleAlaAsnMetIleValGlyAlaThrPheLeuThr 407
Qy 1419 GTGTCGAAGCCATCACATCGCGGAGGTTCTCTCTACTCCGGCATCGCGGCTC 1478
Db 408 MetLeuAsnThrLeuGlyAsnAlaAsnThrPheTrpValTrpAlaAlaLeuAsnValLeu 427
Qy 1479 GCCTGGGTGTTCTTCTACACCTACTCCCGGAGACCGCGCGGCGGAGGATG 1538
Db 428 PheLeuLeuLeuThrLeuTrpLeuValProGluThrLysHisValSerLeuGluHisIle 447
Qy 1539 AGCAAG 1544
Db 448 GluArg 449

RESULT 10
ID ITR2_YEAST STANDARD; PRT; 612 AA.
AC P30606;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myo-inositol transporter 2.
GN ITR2 OR YOL103W OR HRB612.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91250431; PubMed=2040626;
RA Nikawa J.-I., Tsukagoshi Y., Yamashita S.;
RT "Isolation and characterization of two distinct myo-inositol
RT transporter genes of Saccharomyces cerevisiae.";
RL J. Biol. Chem. 266:11184-11191(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96076631; PubMed=7502582;
RA Vandenbol M., Durand P., Portetelle D., Hilger F.;
RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV
RT including the Tyl-H3 retrotransposon, the sufi(+) frameshift
RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a
RT delta element.";
RL Yeast 11:1069-1075(1995).
CC -!- FUNCTION: MINOR TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90353; BAA14367.1; -;
DR EMBL; Z48149; CAA8159.1; -;
DR EMBL; Z74845; CAA9119.1; -;
DR PIR; B40538; B40538.
DR GERMOnline; 143525; -;
DR SGD; S0005463; ITR2.
DR GO; GO:0005365; F:myo-inositol transporter activity; IMP.
DR GO; GO:0015798; P:myo-inositol transport; IMP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005829; Sug transporter.
DR InterPro; IPR003663; Sugar transp.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPRT.

TIGRFAMS; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 107
FT TRANSMEM 108 128
FT DOMAIN 129 155
FT TRANSMEM 156 176
FT DOMAIN 177 182
FT TRANSMEM 183 203
FT DOMAIN 204 212
FT TRANSMEM 213 233
FT DOMAIN 234 241
FT TRANSMEM 242 262
FT DOMAIN 263 272
FT TRANSMEM 273 293
FT DOMAIN 294 362
FT TRANSMEM 363 383
FT DOMAIN 384 402
FT TRANSMEM 403 423
FT DOMAIN 424 426
FT TRANSMEM 427 447
FT DOMAIN 448 457
FT TRANSMEM 458 488
FT DOMAIN 489 512
FT TRANSMEM 513 533
FT DOMAIN 534 536
FT TRANSMEM 537 557
FT DOMAIN 558 612
FT CARBOHYD 337 397
SQ SEQUENCE 612 AA; 67041 MW; 48E733102EAD7298 CRC64;

Alignment Scores:
Pred. No.: 1,48e-20 Length: 612
Score: 572.50 Matches: 154
Percent Similarity: 50.0% Conservative: 103
Best Local Similarity: 29.9% Mismatches: 210
Query Match: 15.8% Indels: 47
DB: 14 Gaps: 14

US-10-051-902A-21 (1-2017) x ITR2_YEAST (1-612)
Qy 117 GAGGCGTCGCGCCGAGAGAGAGGCAAGTCGCTTCGCTTCGCTCGCCATCCTC 176
Db 106 GlnSerIleSerPro-----PheIleIleThrLeuThrPheVal 118
Qy 177 GCCTCCATGACCTCCATCTCTCGGTCAGCATATCGGGGTGATGAGCGGCGCTG 236
Db 119 AlaSerIleSerGlyPheMetPheGlyTyraThrGlyTyraSerSerAlaLeuIle 138
Qy 237 TACATCAAGAGGACTTCAAC-----ATCAGTACGCGAAGAGGTGAGGTCTCATG 287
Db 139 SerIleAsnArgAspLeuAspAsnLysValLeuThrTyraGlyGluLysGluLeuIleThr 158
Qy 288 GGCATACCTGAACCTTACTCTGCTCATGCTTCCTCGCGGCGGCGGAGCTCGGACTG 347
Db 159 AlaAlaThrSerLeuGlyAlaLeuIleThrSerValGlyAlaGlyThrAlaAlaAspVal 178
Qy 348 ATCGSCCGCGGTACACCATCTGTTCGCCCGCTCATATTCTTCGCGGGSGGTTCCTC 407
Db 179 PheGlyArgArgProCysLeuMetPheSerAsnLeuMetPheLeuIleGlyAlaIleLeu 198
Qy 408 ATGGGTTTCGCGGTCACTACGCTCATGTCGCGCTTCGCGCGCTTCGTCGCGGATCGGC 467
Db 199 GlnIleThrAlaHisLysPheTrpGlnMetAlaAlaGlyArgLeuIleMetGlyPheGly 218
Qy 468 GTGGGCTACGCGCTCATGATCGCGCGGTGACACCGCGAGGTGTGCGCGCGCTCGCGC 527
Db 219 ValGlyIleGlySerLeuIleSerProLeuPheIleSerGluIleAlaProLysMetIle 238
Qy 528 CGTGCTTCCTGAGCTGCTTCGCGAGGTGTTCATACACTTCGCGCATCTCTCTCGGGTAC 587
Db 528 CGTGCTTCCTGAGCTGCTTCGCGAGGTGTTCATACACTTCGCGCATCTCTCTCGGGTAC 587

DR InterPro; IPR005829; Sug transporter.
 DR InterPro; IPR003663; Sugar transpt.
 DR Pfam; PF00083; sugar tr; 1-
 DR PRINTS; PR00171; SUGRTNSPT.
 DR TIGRFAMs; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR TRANSPORT 1; FALSE_NEG.
 DR PROSITE; PS00217; SUGAR TRANSPORT 2; 1.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT DOMAIN 1 29
 FT TRANSMEM 30 50
 FT DOMAIN 51 59
 FT TRANSMEM 60 80
 FT DOMAIN 81 92
 FT TRANSMEM 93 113
 FT DOMAIN 114 120
 FT TRANSMEM 121 141
 FT DOMAIN 142 155
 FT TRANSMEM 156 176
 FT DOMAIN 177 184
 FT TRANSMEM 185 205
 FT DOMAIN 206 263
 FT TRANSMEM 264 284
 FT DOMAIN 285 301
 FT TRANSMEM 302 322
 FT DOMAIN 323 331
 FT TRANSMEM 332 352
 FT DOMAIN 353 373
 FT TRANSMEM 374 400
 FT DOMAIN 401 421
 FT TRANSMEM 422 423
 FT DOMAIN 424 444
 FT TRANSMEM 445 482
 FT DOMAIN 482 482
 SQ SEQUENCE 482 AA; 52756 MW; 90851C4F4C48EE01 CRC64;

Alignment Scores:
 Pred. No.: 9,49e-20 Length: 482
 Score: 555.50 Matches: 152
 Percent Similarity: 51.92% Conservative: 105
 Best Local Similarity: 30.71% Mismatches: 195
 Query Match: 15.37% Indels: 44
 DB: 1 Gaps: 13

US-10-051-902A-21 (1-2017) x YFIF_BACSU (1-482)

QY 117 GAGCGCTCGCGCGAAGAG-----AAGGCAACGTCGGTTCGGCTTCGCC 164
 Db 7 GluAlaValIleGlyLysGlySerLeuAlaHisLysGlyLeuLeuArg-----Thr 23
 QY 165 TGGCCATCTCGCTCCATGACCTCCATCTCTCGGCTACCATATCGGGGTGATGAGC 224
 Db 24 IleThrLeuValSerThrPheGlyGlyLeuPheGlyTyrAspThrGlyValIleAsn 43
 QY 225 GGGCGCTGCTGATCAACAAG-----GACTTCAACATCATGACGGGAAGTGGAG 278
 Db 44 GlyAlaLeuProPheMetAlaThrAlaGlyGlnLeuAsnLeuThrProValThrGluGly 63
 QY 279 GTTCTCATGGCATACTGACCTACTCGCTCATCGGCTCCTTCGGCGGGGGGAGC 338
 Db 64 LeuValAlaSerSerLeuLeuLeuGlyAlaAlaPheGlyAlaMetPheGlyGlyArgLeu 83
 QY 339 TCGAAGTGGATCGCGCGGTACACCATCGTGTTCGCCCGCTCATATTCTTCGGCGGG 398
 Db 84 SerAspArgHisGlyArgGlyThrIleLeuTyrLeuAlaLeuPheIleAlaAla 103
 QY 399 GSGTTCCTCATGGGTTTCGGGTGCTCACTAGCCATGCTCATGTTCGGCGCTTCGGGCC 458
 Db 104 ThrLeuGlyCysThrPheSerProAsnAlaSerValMetIleAlaPheArgPheLeu 123
 QY 459 GGCATCGGCTGGGCTACGGCTCATGATCGCGCGGTGTACACCGCCGAGGTGTCGCG 518
 Db 124 GlyLeuAlaValGlyCysAlaSerValThrProThrPheLeuAlaGluIleSerPro 143

QY 519 GGGTCGGCGCGTTCCTGACGCTTCCTCCGGAGGTGTTTCATCAACTTCGGCATCCG 578
 Db 144 AlaGluArgArgGlyArgIleValThrGlnAsnGluLeuMetIleValIleGlyGlnLeu 163
 QY 579 CTCGGGTACGTCGGAACATATGCTTTCTCCCGCTTCGCGCTGAACCTCGG----- 629
 Db 164 LeuAlaTyrThrPheAsn-----AlaIleIleGlySerThrMetGlyGluSerAla 180
 QY 630 -----TGGCGCATCATGCTCGGCATCGCGCGCGCGCTCGCTCGCTCGCTCGCTCATG 683
 Db 181 AsnValTrpArgTyrMetLeuValIleAlaThrLeuProAlaValLeuTrpPheGly 200
 QY 684 GTCTCGGCATCCCGAGTCGCGCGGTGGTGGTGCATCAAGGGAGCGCTCGCGACGCC 743
 Db 201 MetLeuIleValProGluSerProArgTrpLeuAlaAlaLysGlyArgMetGlyAspAla 220
 QY 744 AAGGTGGTCTCGAGAACCTCCGACACGGCGGAGGAGCGCGGAGCGCTGCGCGAC 803
 Db 221 LeuArgValLeuArgGlnIleArgGluAspSer---GlnAlaGlnGlnIleLysGlu 239
 QY 804 ATCAAGCGCGCGCGCATCCCTGAGGAGTTCGACGGCGAGCTGTCACCTCCCGAACG 863
 Db 240 IleLysHisAla-----IleGluGly-----ThrAlaLysLys 250
 QY 864 AGAGGAGCGGAAACGAGAGCGGGTGTGGAAGGAGCTCATCTCTCCCGACCCCGCGC 923
 Db 251 AlaGlyPheHisAspPheGlnGluProTrp----- 260
 QY 924 ATCGCGCGCATCTCTGTCGCGGATCGGCATCCACTTCTTCCAGCATCGCTGGGCAAT 983
 Db 261 IleArgArgIleLeuPheIleGlyIleAlaIleValGlnGlnIleThrGlyVal 280
 QY 984 CACTCGTCTGCTTCTCAGACGCTCTGCTGTTCGAAGCGCGCGGATTAACGACACAAA 1043
 Db 281 AsnSerIleMetTyrTyrGlyThrGluIleLeuArgGluAlaGlyPheGlnThrGluAla 300
 QY 1044 CACTTCTTGGCACCATCTGGCGTTCGCTGCACCAAGAGCGCTTTCATCTGTGTGGC 1103
 Db 301 AlaLeuIleGlyAsnIle---AlaAsnGlyValIleSerValIleAlaValIlePheGly 319
 QY 1104 ACTTCTTTCATCGACGGCTCGCGCGCGCGCTGTGCTGGGACGACGCGCGGGATA 1163
 Db 320 IleTrpLeuLeuGlyLysValArgArgProMetLeuIleIleGlyGlnIleGlyThr 339
 QY 1164 ATCTCTCTCTCATCGGCTCGCGCGCGGCTCACCGTCGTCGGCGACGACCGCGCGC 1223
 Db 340 MetThrAlaLeuLeuIleGlyIle---LeuSerIleValLeuGluGlyThrProAla 358
 QY 1224 AAGATACCTTGGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1283
 Db 359 ---LeuProTyrValVal-----LeuSerLeuThrIleLeuPheLeuAlaPheGln 374
 QY 1284 TCATCGGCTTGGCGCGCATCACGTCGGTGTACAGCTCGAGATCTTCCCGCTCCAGGTG 1343
 Db 375 GlnThrAlaIleSerThrValThrTrpLeuMetLeuSerGluIlePheProMetHisVal 394
 QY 1344 CGCGCGCTGGGCTCGCTCGCTCGGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1403
 Db 395 ArgGlyLeuGlyMetGlyIleSerThrPheCysLeuTrpThrAlaAsnPheLeuIleGly 414
 QY 1404 ATGACTTCTCTGCTCGCTTCCAAAGCGCATCACCATCGCGCGCGCGCTTCTCTACTCC 1463
 Db 415 PheThrPheProIleLeuLeuAsnHisIleGlyMetSerAlaThrPheIlePheVal 434
 QY 1464 GGCATCGCGCGCTCGCTCGGCTGTCTTCTACACCTACCTCCCGGAGACCGCGCGCGG 1523
 Db 435 AlaMetAsnIleLeuAlaIleLeuPheValLysLysTyrValProGluThrLysGlyArg 454
 QY 1524 ACGTGGAGGAGAGT-GAGCAAGCTGTTCGGCGACACGCGCGCGCGCGCGCGCG 1567
 Db 455 SerLeuGluGlnLeuGluHisSerPheArgGlnTyrGlyArgArg 469

RESULT 12

Db 121 sAlaGlyAlaLeuLeuLeuLeuGlyPheAlaAlaLeuHisValTrpMetLeuLeuValGlyArgI 141
QY 452 CGTGGCCGATCGCGCGTGGCTACGGCTCATCGCTCATCGCGCTGACACCGCGCGAGT 511
Db 141 eLeuLeuGlyPheGlyIleGlyPheAlaAsnGlnAlaValProLeuLeuTrpLeuSerGluMe 161
QY 512 GTCCCGCGCTCGCGCGTGGCTTCCGACGCTGTTCCCGAGGAGTGTTCATCAACTCGG 571
Db 161 tAlaProTrpLysTrpArgGlyAlaLeuAsnIleGlyPheGlnLeuSerIleThrIleG 181
QY 572 CATCTGCTCGGCTACGCTCGCAACTATGCTTTCCTCCGCTTCCGCTGCGCTGAACCTCGCGTG 631
Db 181 yLeuLeuValAlaGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 201
QY 632 CGCATCATGCTCGCATCGCGCGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 691
Db 201 pArgLeuSerLeuGlyAlaValAlaValProAlaLeuIleIleThrIleGlySerLeuVa 221
QY 692 CATCGCGAGTCGCGCGTGGCTCATGAAGGAGCGCTCGCGACGCCAAGGTGT 751
Db 221 lLeuProAspThrProAsnSerMetIleGluArgGlyGln----- 234
QY 752 GCTGAGAAGACCTCCGACACGCGGAGGAGCGCGGAGCGCTGCGCGACATCAAGGC 811
Db 235 -----HisGluGluAlaLeuThrLysLeuArgIleArgG 247
QY 812 CGCGCGCGATCCTGAGGAGCTCGACGCGAGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 871
Db 247 yValAspValSerGlnGluPheAsp---AspLeuValAlaAlaSerLysGluSerG 266
QY 872 CGAAACGAGAGCGGTGGAGAGAGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 931
Db 266 nSerIleGlu---HisProTrpArgAsnLeuLeu-----ArgArgLysTrpArgPr 282
QY 932 CATCTGCTGCTGCGGATCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991
Db 282 cHisLeuThrMetAlaValMetIleProPheGlnGlnLeuThrGlyIleAsnValI 302
QY 992 CGTCTTCTACAGCCCTCTCGTGTTCAGAGCGCGGATTAAACGACGACAAACACTTCT 1051
Db 302 eMetPheTyAlaProValLeuPheAsnThrIleGlyPheThrThrAspLaserLeuMe 322
QY 1052 GGGCACCACATGGCGCTCGGTGTCACCAAGAGCTTTTCATC-----TGTGGC 1102
Db 322 tSerAlaVal-----ValThrGlySerValAsnValAlaAlaThrLeuValSe 338
QY 1103 GACTTTCTTCATCGCGCTCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1159
Db 338 rIleTyGlyValAspArgTrpGlyArgArgPheLeuPheLeuGluGlyGlyThrGlnMe 358
QY 1160 GATAATCTCTCCCTCATC-----GGCTCGCGCGCGCTCACCGCTCGCTCGGCCA 1210
Db 358 tLeuIleCysGlnAlaValAlaAlaAlaCysIleGlyAlaLysPheGlyValAspGlyTh 378
QY 1211 GCACCCGCGCAGATACCTTGGGCGCATCGGCTTGAAGCATCGCTCCACCTCCGCTA 1270
Db 378 r---ProGlyLeuProLysTrpTrpAlaIleValValValThrPheIleCysIleTy 397
QY 1271 CGTGGCTTCTTCTCCATCGCTTGGCCCATCACGTGGGTGTACAGCTCGGAGATCTT 1330
Db 397 rValAlaGlyPheAlaTrpSerTrpGlyProLeuGlyTrpLeuValProSerGluLeuPh 417
QY 1331 CCCTCGAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1390
Db 417 eProLeuGluIleArgSerAlaAlaGlnSerIleThrValSerValAsnMetIlePheTh 437
QY 1391 CGGCTCATCTCATGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1450
Db 437 rPheIleAlaGlnIlePheLeuThrMetLeuCysHisLeuLysPhe---GlyLeuPh 456
QY 1451 CTTCCTTACTCGCGCATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1510
Db 456 eLeuValPheAlaPhePheValValValMetSerIlePheValTyIlePheLeuProG 476

QY 1511 GACCCGCGCGCGACGCTGAGGAGATGAGCAAGCTGTC 1550
Db 476 uThrLysGlyIleProIleGluMetGlyGlnValTrp 489

RESULT 13

HUP3_CHLKE
ID HUP3_CHLKE STANDARD; PRT; 534 AA.
AC Q39525;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE H(+)/hexose cotransporter 3.
GN HUP3.
OS Chlorella kessleri.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellales incertae sedis; Parachlorella.
OC NCBI_TaxID=3074;
RN [1]_TaxID=3074;
RP SEQUENCE FROM N.A.
RX MEDLINE=9517527; PubMed=7870840;
RA Stadler R., Wolf K., Hilgarth C., Tanner W., Sauer N.K.;
RT "Subcellular localization of the inducible Chlorella HUP1
monosaccharide-H⁺ symporter and cloning of a Co-induced galactose-H⁺
symporter";
RT symporter";
RL Plant Physiol. 107:33-41(1995).
CC -!- FUNCTION: Active uptake of hexoses.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL, X75440; CA53192.1; -
PIR; S38435; S38435.
InterPro; IPR007114; MFS.
InterPro; IPR005828; Sub transporter.
InterPro; IPR005829; Sug transporter.
InterPro; IPR003663; Sugar transp.
Pfam; PF00083; sugar tr; 1-
PRINTS; PR00171; SUGTRANSPORT.
TIGRfams; TIGR00879; Sp; 1.
PROSITE; PS00850; MFS; 1.
PROSITE; PS00216; SUGAR TRANSPORT 1; 1.
PROSITE; PS00217; SUGAR TRANSPORT 2; 1.
KW Transmembrane; Transport; Sugar transport; Symport; Repeat;
KW Multigene family.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 41 1 (POTENTIAL).
FT DOMAIN 42 86 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 87 107 2 (POTENTIAL).
FT DOMAIN 108 113 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 114 135 3 (POTENTIAL).
FT DOMAIN 136 137 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 138 158 4 (POTENTIAL).
FT DOMAIN 159 173 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 174 194 5 (POTENTIAL).
FT DOMAIN 195 204 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 205 225 6 (POTENTIAL).
FT DOMAIN 226 299 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 300 322 7 (POTENTIAL).
FT DOMAIN 323 328 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 329 349 8 (POTENTIAL).
FT DOMAIN 350 352 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 353 373 9 (POTENTIAL).
FT DOMAIN 374 394 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 395 415 10 (POTENTIAL).
FT DOMAIN 416 433 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 434 454 11 (POTENTIAL)
 FT DOMAIN 455 456 EXTRACELLULAR (POTENTIAL)
 FT TRANSMEM 457 477 12 (POTENTIAL)
 FT DOMAIN 478 534 CYTOPLASMIC (POTENTIAL)
 SQ SEQUENCE 534 AA; 57772 MW; 52E03715F3726017 CRC64;

Alignment Scores:
 Pred. No.: 5,86e-19 Length: 534
 Score: 539.00 Matches: 155
 Percent Similarity: 47.70% Conservative: 83
 Best Local Similarity: 31.06% Mismatches: 205
 Query Match: 14.91% Indels: 56
 DB: 1 Gaps: 16

US-10-051-902a-21 (1-2017) x HUP3_CHLXE (1-534)

QY 141 GGCAAGTCGCGTTCGCTTCGCTCGCGCAATCTCGCTCCATGACCTCCATCTCTCTC 200
 Db 21 GlyLeuThrAlaTyrValLeuValAlaLeuValAlaAlaCysGlyMetLeu 40
 QY 201 GGCTACGATATCGGGTCATGACGGCGGCGTCG----- 233
 Db 41 GlyTyrAspAsnGlyValThrGlyGlyValAlaSerMetGluInPheGluArgLysPhe 60
 QY 234 -----CTGTACATCAAGAAGGAC-----TTCAACATC 260
 Db 61 PheProAspValTyrGluLysLysGlnInleValGluThrSerProTyrCysThrTyr 80
 QY 261 AGTAGCGGAGAGTGAGGTCTCATGGCATATGACCTCTACTCGCTCATCGGCTCC 320
 Db 81 AspAsnProLysLeuGlnLeuPheValSerLeuPheLeuAlaGlyLeuSerCys 100
 QY 321 ---TTCCGGCGGGCGGACCTCGATCGGCTGATCGCGCGGCGGCTACACATCGTTCGCC 377
 Db 101 IlePheSerAlaTyrPheThrArgAsnTrp---GlyArgLysAlaSerMetGlyLeGly 119
 QY 378 GCGTCATATCTTC---GCGGGGGGTTCCTCATGGGGTTCGCGCTCACTACGCCATG 434
 Db 120 GlyIlePhePheIleAlaAlaGlyLeuValAsnAlaPheAlaGlnAspIleAlaMet 139
 QY 435 CTCATGTTCCGCGCTTCGTGCGCGCATCGGCTGCGCTACGGCTCATGATCGCGCGG 494
 Db 140 LeuIleValGlyArgValLeuLeuGlyPheGlyValGlyLeuGlySerGlnValValpro 159
 QY 495 GTGTACACCGCGAGGTCTCCGCGCTCGCGGCTTCCTGACGTCTTCCTCCGAG 554
 Db 160 GlnTyrLeuSerGluValAlaProPheSerHisArgGlyMetLeuAsnIleGlyTyrGln 179
 QY 555 GTGTCATCAACTCGCGCATCTCTCGGTCAGCTCTCGAATATGCTTCTCCGCTTG 614
 Db 180 LeuPheValThrIleGlyIleLeuIleAlaGlyLeuValAsnTyrGlyValArgAsn--- 198
 QY 615 CGCTGAACCTCGGTGCGCATGTCTCGGCATCGCGCGGCGGCTCGCTCTCTCTC 674
 Db 199 ---TrpAspAsnGlyTyrArgLeuSerLeuGlyLeuAlaAlaValProGlyLeuIleLeu 217
 QY 675 GCGCTCATGTCTCGGATCGCGAGTCGCGGCTGGTGTGTCATGAGGCGCTC 734
 Db 218 LeuLeuGlyAlaIleValLeuProGluSerProAsnPheLeuValGluLysGlyArgThr 237
 QY 735 GGGAGCGCAAGTGTGTGTGAGAAAGACCTCCGACAGCGGAGGAGCGCGGAGCGG 794
 Db 238 AspGlnGlyArgArgIleLeuGluLysLeuArgGlyThrSerHisValGluAlaGlu--- 256
 QY 795 CTGGCGCATCAAGCGCGCGCGCATCTCCTGAGGAGCTCGAGCGGCGAGTGGTAC 854
 Db 257 PheAlaAspIleValAlaAlaValGluIleAlaArgProIle----- 270
 QY 855 GTCCCCAAGAGGAGCGGAAACAGAGAGGGGTGTGGAAGGAGCTCATCTCTCCCGG 914
 Db 271 -----ThrMetArgGlnSerTrpA-gSerLeuPhe----- 280
 QY 915 ACCCGCGCGCATCGCGGATCTGTCTCGGATTCGGCATCCACTCTTCACGATCGG 974

Db 281 ThrArgGlyTyrMetProGlnLeuLeuThrSerPheValIleGlnPheGlnPhe 300
 QY 975 TTGGGCACTACTCGTGTCTTCTACAGCCCTCTCGTGTTCAGAGCCCGGATTAACG 1034
 Db 301 ThrGlyIleAsnAlaIlePheTyrValProValLeuPheSerSerLeuGlySerAla 320
 QY 1035 AACGACAAACACTTCTTGGGACCACTTGGCGCTTCGCGTTCACCAAGAGGCTTTTCATC 1094
 Db 321 SerSerAlaAlaLeuLeuAsnThrVal---ValValGlyAlaValAsnValGlySerThr 339
 QY 1095 TTGTTGGGCACTTTCTTATCAGCGGCTCGGCGCGCGCTTCTCTGGGACGACG 1154
 Db 340 MetIleAlaValLeuLeuSerAspLysPheGlyArgArgPheLeuLeuIle-----Glu 357
 QY 1155 GCGGGATAATCTCTCTCTCATCGGCTCGGCGCGGCTCACCGTC-----GTC 1205
 Db 358 GlyGlyIleThrCysCysLeuAlaMetLeuAlaAlaGlyIleThrLeuGlyValGluPhe 377
 QY 1206 GCGGACACCGCGACCAAGATACCTTGGGCGCTTGGGCGCTTAAGCATCGCTCCACCCCTC 1265
 Db 378 GlyGlnTyrGlyThrGluAspLeuProHisProValSerAlaGlyValLeuAlaValIle 397
 QY 1266 GCC---TAGTCGCGCTTCTTCTCCATCGGCTTGGCGCGCTCACCGTGTACAGCTCG 1322
 Db 398 CysIlePheIleAlaGlyPheAlaTrpSerTrpGlyProMetGlyTrpLeuIleProSer 417
 QY 1323 GAGATCTTCCCGCTCAGGTGCGCGCTGCTGCTCGCGGCTCGCGCGCAACCGC 1382
 Db 418 GluIlePheThrLeuGluThrArgProAlaGlyThrAlaValAlaValMetGlyAsnPhe 437
 QY 1383 GTACACAGCGGCTCATCTCCATGACCTTCTGCTGCTCAAGCGCATCACCATCGGC 1442
 Db 438 LeuPheSerPheValIleGlyGlnAlaPheValSerMetLeuCysAlaMetLysPhe--- 456
 QY 1443 GCGAGCTTCTCTCTACTCTCCGCG-----ATCGCGCGCTCGCTCGGTGTCTTC 1493
 Db 457 GlyValPheLeuPhePheAlaGlyTrpLeuValIleMetValLeuCysAlaIlePheLeu 476
 QY 1494 TACACCTACTCTCGGAGACCGCGCGCGGCTGAGGAGATGAGCAAGCTGTTC 1550
 Db 477 -----LeuProGluThrLysGlyValProIleGluArgValGlnAlaLeuTyr 492

RESULT 14
 HUP1_CHLXE STANDARD; PRT; 534 AA.
 ID HUP1_CHLXE
 AC P15686; Q39523;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE H(+)/hexose cotransporter 1.
 GN HUP1.
 OS Chlorella kessleri.
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
 OC Chlorellales incertae sedis; Parachlorella.
 OC NCBI_TaxID=3074;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90092536; PubMed=2599110;
 RA Sauer N., Tanner W.;
 RT "The hexose carrier from Chlorella. cDNA cloning of a eucaryotic H+-
 cotransporter.";
 RL FEBS Lett. 259:43-46 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91330335; PubMed=1868571;
 RA Wolf K., Tanner W., Sauer N.;
 RT "The Chlorella H+/hexose cotransporter gene.";
 RL Curr. Genet. 19:215-219 (1991).
 CC -!- FUNCTION: Active uptake of hexoses.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- INDUCTION: BY glucose.
 CC -!- SIMILARITY: Belongs to the sugar transporter family.

 This SWISS-PROT entry is copyright © It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL databanks at
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

[illegible]

US-10-051-902A-21 (1-2017) x HUP1 CHLKE (1-534)

[illegible]


```
QY 414 TTCGCCGTCAACTAGCCATGCTCATGTTCCGGCGCTTCGTGGCCGGCATCGCGCGTGGGC 473
Db 103 LeuProPheThrIleTrpAspPheIlePheTrpArgValLeuGlyGlyIleGlyValGly 122
QY 474 TACCGCTCATGATCGCGCGGTGTACACCGCGAGGTGTCCGGCGGTCCGGCGGTGC 533
Db 123 AlaAlaSerValIleAlaProAlaTrpIleAlaGluValSerProAlaHisLeuArgGly 142
QY 534 TTCTGTAGCTGTTCCCGAGGTGTTTCATCACTTCGGCATCTTCGTGGGTACGTCTCG 593
Db 143 ArgLeuGlySerLeuGlnLeuAlaIleValSerGlyIlePheIleAlaLeuLeuSer 162
QY 594 AACTATGCTTTCCTCCCGCTTG-----CCGCTGAACCTCGGG--- 629
Db 163 AsnTrpPheIleAlaLeuMetAlaGlyGlySerAlaGlnAsnProTrpLeuPheGlyAla 182
QY 630 -----TGGCGCATCATGCTCGGCATCGCGCGCGCTCGCTGCTCGCGCTCATG 683
Db 183 AlaAlaTrpArgTrpMetPheTrpThrGluLeuIleProAlaLeuLeuTyrGlyValCys 202
QY 684 GTGCTCGCATGCGCGAGTCCGCGGTGCTGCTCATGAAGAGCGCCTCGCGGACGCC 743
Db 203 AlaPheLeuIleProGluSerProArgTyrLeuValAlaGlnGlyGlyGluLysAla 222
QY 744 AAGTGTGTGCTGAGAGAGACTCCGACACCGCGGAGAGCGCGGAGCGCTGCGCGAC 803
Db 223 AlaAlaIleLeuTrpLysValGlu-----GlyGlyAspValProSerArgIleGluGlu 240
QY 804 ATCAAGCGCCCGCGCATCCCTAGGAGCTCGACGGCGACGTGTGACCGTCCCGCAAG 863
Db 241 IleGlnAlaThrValSerLeuAspHisLys-----ProArg 252
QY 864 AGAGGAGCGGAACGAGAGCGGTGTGAAGAGCTCATCTGTCCCGACCCCGGCC 923
Db 253 PheSerAsp-----LeuLeuSerArgArgGlyGly 262
QY 924 ATCGCGGCATCTCTGTCGGGATCGGCATCGCATCTTCTCCAGCATGCTTGGGCATT 983
Db 263 LeuLeuProIleValTrpIleGlyMetGlyLeuSerAlaLeuGlnGlnPheValGlyIle 282
QY 984 CACTCGCTCTTCTACAGCCTCTCGTGTTCAGAGCCCGGATTAAAGACAGACAAA 1043
Db 283 AsnValIlePheTyrTyrSerSerValLeuTrpArgSerValGlyPheThrGluGluLys 302
QY 1044 CACTTCTTGGGACCACTTGGCGGTTCGTGTCCACCAAGAGGCTTTCATCTGTGGCG 1103
Db 303 SerLeuLeuIleThrValIleThr---GlyPheIleAsnIleLeuThrThrLeuValAla 321
QY 1104 ACTTCTTTCATCGACGCGCTCGGCGCGCGCTGTGTGTG---GGCAGCAGCGGC--- 1157
Db 322 IleAlaPheValAspLysPheGlyArgLysProLeuLeuMetGlySerIleGlyMet 341
QY 1158 -----GGGATAATCTCTCTCCCTCATCGGCTCGGCGCGGCTCACCGTCGTC 1205
Db 342 ThrIleThrLeuGlyIleLeuSerValValPheGly-----GlyAlaThrValVal 358
QY 1206 GGCCAGCACCCGAGCAAGATACCTTGGCCCATCGGC---CTAAGCATCGCTCCACC 1262
Db 359 AsnGlyGlnPro-----ThrLeuThrGlyAlaAlaGlyIleIleAlaLeuValThrAla 376
QY 1263 CTGCGCTACGTGCGCTTCTTCTCCATCGGCTTGGCCCATCATCGTGGGTGTACAGCTCG 1322
Db 377 AsnLeuTyrValPheSerPheGlyPheSerTrpGlyProIleValTrpValLeuLeuGly 396
QY 1323 GAGATCTTCCCGTCCAGGTGCGCGCTGCTGCTCGCTCGGCTCGCGCGCAACCGC 1382
Db 397 GluMetPheAsnAsnLysIleArgAlaAlaLeuSerValAlaAlaGlyValGlnTrp 416
QY 1383 GTCAACAGCGCGTCACTCTCCATGACCTTCTGCTGCTGTCTCAAGGCCCATCACCATCGGC 1442
Db 417 IleAlaAsnPheIleIleSerThrThrPheProLeuLeuAspThrValGlyLeuGly 436
QY 1443 GGCAGCTTCTTCTCTACTCCGCGCATCGCGCGCTCGCGTGGGTGTCTTCTACACCTAC 1502
```

```
Db 437 ProAlaTyrGlyLeuTyrAlaThrSerAlaAlaIleSerIlePhePheIleTrpPhePhe 456
QY 1503 CTCCTCCGAGACCCGCGCGCGGACGCTGGAGGAGATG 1538
Db 457 ValLysGluThrLysGlyThrLeuGluGlnMet 468
```

Search completed: June 30, 2004, 18:44:40
Job time : 71.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:38:20 ; Search time 103.5 Seconds

(without alignments)
12297.589 Million cell updates/sec

Title: US-10-051-902A-21

Perfect score: 3614

Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaaaa 2017

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0			
Xgapext 0.5			
Fgapop 6.0			
Delop 6.0			

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO.spool.P/US10051902/tunat 30062004 164723 20411/app.query.fasta_1.2183
-DB=SPRMBL 25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051902@cgn 1_112@runat 30062004 164723 20411 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPRMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mmc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_virus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2560	70.8	577	10 Q945E5	Q945E5 oryza sativ

ALIGNMENTS

RESULT 1

Q945E5 ID Q945E5 PRELIMINARY; PRT; 577 AA.

AC Q945E5; DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative sugar transporter.
OS Oryza sativa (Rice)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.

RA Wu P., Chen Q., Huang G., Yi K.;

RT Molecular cloning of putative sugar transporter in rice.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

DR EMBL; AF416867; AAL14615.1; -

DR Gramene; Q945E5; -

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005351; F:sugar porter activity; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

	511	511	10	Q8LHC3
2	1742	48.2	511	10 Q8LHC3
3	1710	47.3	538	10 Q84QH3
4	1625.5	45.0	539	10 Q9LS92
5	1817	44.7	523	10 Q7XA50
6	1607	44.5	509	10 Q84KI7
7	1596.5	44.2	481	10 Q84RI1
8	1590	44.0	545	10 P93076
9	1579	43.7	511	10 Q9XIH7
10	1559	43.1	511	10 Q9XIH6
11	1519	42.0	491	10 Q84RI2
12	1513	41.9	513	10 Q9FQX3
13	1494	41.3	519	10 Q7XB36
14	1398	38.7	523	10 Q8RVQ2
15	1385.5	38.3	508	10 Q9ZN80
16	1361	37.7	493	10 Q23213
17	1307.5	36.2	574	10 Q9AUM9
18	1304	36.1	506	10 Q8W2K8
19	1186	32.8	538	10 Q7XKF1
20	1167.5	32.3	535	10 Q7XKF0
21	1143	31.6	547	10 Q9SKT9
22	1134.5	31.4	523	10 Q7X6M3
23	673.5	18.6	521	10 Q22848
24	670.5	18.6	457	10 P96742
25	669.5	18.5	509	10 Q8VZK6
26	653	18.1	467	10 Q8CQA7
27	649.5	18.0	596	10 Q7X120
28	643.5	17.8	498	10 Q84UY4
29	639.5	17.7	618	11 Q921A2
30	630	17.4	245	10 Q8VZ80
31	629.5	17.4	582	10 Q23492
32	626	17.3	580	10 Q9C757
33	623	17.2	493	10 Q7XRI6
34	618	17.1	581	10 Q9LKH1
35	612	16.9	409	16 Q83EH4
36	609.5	16.9	473	16 Q34718
37	608.5	16.8	544	10 Q93WT7
38	608	16.8	558	10 Q9FIF2
39	600.5	16.6	525	10 Q852B0
40	594	16.4	517	16 Q8G3X1
41	591.5	16.4	546	10 Q93241
42	591	16.4	581	10 Q7XQ00
43	589.5	16.3	479	2 Q9FDM0

Q8LHC3 oryza sativ
Q84QH3 prunus cera
Q9LS92 arabidopsis
Q7XA50 glycine max
Q84KI7 prunus cera
Q84RI1 malus domes
P93076 beta vulgar
Q9XIH7 arabidopsis
Q9XIH6 arabidopsis
Q84RI2 malus domes
Q9FQX3 apium grave
Q7XB36 orobanche r
Q8RVQ2 apium grave
Q9ZN80 arabidopsis
Q23213 arabidopsis
Q9AUM9 oryza sativ
Q8W2K8 oryza sativ
Q7XKF1 oryza sativ
Q7XKF0 oryza sativ
Q9SKT9 arabidopsis
Q7X6M3 oryza sativ
Q22848 arabidopsis
P96742 bacillus su
Q8VZK6 arabidopsis
Q8CQA7 staphylococ
Q7X120 oryza sativ
Q84UY4 mesembryant
Q921A2 rattus norv
Q8VZ80 arabidopsis
Q23492 arabidopsis
Q9C757 arabidopsis
Q7XRI6 oryza sativ
Q9LKH1 mesembryant
Q83EH4 coxiella bu
Q34718 bacillus su
Q93WT7 olea europ
Q9FIF2 arabidopsis
Q852B0 oryza sativ
Q8G3X1 bifidobacte
Q93241 arabidopsis
Q7XQ00 oryza sativ
Q9FDM0 zymomonas m

QY	750	GTCTGTGAGNAGACCTCCGACACGGCGGAGGAGCGCGGAGCGCTTGGCGGACATCAAG	809
Db	199	ValLeuGlyIuThrSerAspThrAlaGluGluAlaAlaThrArgLeuAlaGluIleArg	218
QY	810	GCCTCCCGCGGCATCCCTGAGGAGCTCGACGGCGGACGTGTGTGACCGTCCCCACGAGAGGG	869
Db	219	GluAlaValAlaIleProAlaAspLeuAspGlyAspValValaValProLysArg--	237
QY	870	AGCGGAAACAGAAAGCGGGTGTGAAGGAGCTCATCTGTCCCGACCCCGGCATCGGG	929
Db	238	AlaGlyGlyGluArgValTrpLysGluLeuIleLeuSerProThrProAlaValArg	257
QY	930	CGCATCTGTGTCGCGGATCGGCATCCACTTCTTCCAGCATGCTGTTGGCATTCACCTCC	989
Db	258	ArgValLeuLeuSerAlaLeuGlyIleHisPhePheGlnInSerSerGlyIleAspAla	277
QY	990	GTCTGTCTTACAGCCCTCTCTGTTCAGAGCCCGGATTAAACGAAACGACAAACACTTC	1049
Db	278	ValValLeuTyrSerProArgValPheGlnSerAlaGlyIleThrAspLysAsnLysLeu	299
QY	1050	TTGGGCACCACTTGGCGGTTTCGGGTGCACCAAGAGGTTTTTCATCTTGTGGGACCTTC	1109
Db	298	LeuGlyThrThrCysAlaValGlyValThrLysThrLeuPheIleLeuValAlaThrPhe	317
QY	1110	TTCATCCAGCGCTCGGCGCGCGCGCTCTTCTCTGGCAGCAGCGCGGATAATCCTC	1169
Db	318	ThrLeuAspArgPheGlyArgArgProLeuLeuLeuAlaSerAlaGlyGlyMetIleAla	337
QY	1170	TCCCTCATCGGCTCGGCGCGCGGCTCACCGTGTGTGGCCAGACACCCGCGAGCCAAATA	1229
Db	338	ThrLeuValThrLeuGlyLeuGlyLeuThrValIleGlyGluAspAlaThrGlyGly--	356
QY	1230	CCTTGGGCCATCGGCTAAGCATCGCTCCACACCTCGCTACGTCGCTTCTTCTCCATC	1289
Db	357	GlyTrpAlaIleAlaValSerIleAlaSerIleLeuAlaPheValAlaPhePheSerIle	376
QY	1290	GGCTTGGCCCCATCAGTGGGTGTACAGCTCGGAGATCTTCCGCTCCAGGTCCGCGCG	1349
Db	377	GlyLeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuHisLeuArgAla	396
QY	1350	CTGGGTGTCTGCTCGCGCTCGCGCCCAACCGGTACACGCGCGTCATCTCCATGACC	1409
Db	397	LeuGlyCysAlaLeuGlyValGlyLeuAsnArgValThrSerGlyValIleSerMetThr	416
QY	1410	TTCTCTGTCTGCCAAGGCCATCACCATCGCGCGGAGCTTCTTCTCTACTCCGGCATC	1469
Db	417	PheLeuSerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyVal	436
QY	1470	GCGCGCTCCCTGGGTGTTCTTCTACACTACTCCGAGACCCCGCGCGGACACGCTG	1529
Db	437	AlaSerLeuAlaTrpLeuPhePheThrTyrLeuProGluThrArgGlyArgThrLeu	456
QY	1530	GAGGAGATGACGAGCTGTC-----GGCGACACGGCGCGCGCC	1568
Db	457	GluGlnMetGlyGluLeuPheArgIleHisAsnMetAlaGlyAspAspSerAlaAla	476
QY	1569	TCGGATCAGACGAGCGCCAGGAGAGAGAG-----AAGTGGAAATGCGCGCC	1619
Db	477	ThrArgProProSerProGluGluGluGlyIleThrAsnTyrValGluMetAlaAla	496
QY	1620	ACTAACTGATCAAC	1634
Db	497	ProSerSerSerSer	501
RESULT 3			
Q84QH3	ID	PRELIMINARY; PRT; 538 AA.	
Q84QH3	AC		
DT	01-JUN-2003	(T=EMBLrel. 24, Created)	
DT	01-JUN-2003	(T=EMBLrel. 24, Last sequence update)	
DT	01-OCT-2003	(T=EMBLrel. 25, Last annotation update)	
DE		Putative sorbitol transporter.	

GN SORT2.
OS Prunus cerasus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=140311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Montmorency; TISSUE=fruit;
RX MEDLINE=22578918; PubMed=1269316;
RA Gao Z.; Maurouisset L.; Lemoine R.; Yoo S.D.; Van Nocker S.;
RA Loescher W.;
RT "Cloning, Expression, and Characterization of Sorbitol Transporters
RT from Developing Sour Cherry Fruit and Leaf Sink Tissues";
RN Plant Physiol. 131:1566-1575(2003).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Montmorency; TISSUE=fruit;
RA Zhifang G.; Loescher W.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY100638; AAM44082.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr_1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 538 AA; 58569 MW; ADD03B983C6D358 CRC64;

Alignment Scores:
Pred. No.: 5,91e-92 Length: 538
Score: 1710.00 Matches: 341
Percent Similarity: 76.76% Conservative: 62
Best Local Similarity: 64.95% Mismatches: 114
Query Match: 47.32% Indels: 8
DB: 10 Gaps: 4

US-10-051-902a-21 (1-2017) x Q84QH3 (1-538)

QY 81 TCTGACCACCGAGATGGCTTCGCCCGCGTCCGGAGGCGCGTCCGCGGAGGAG 140
Db 13 SerGlyGlnProGlnLysAsnIleAlaAspPheAspProProGlyLysProLysArgAsn 32
QY 141 GGCACAGTCGGTTCGCTTCGCTGGCCATCTCCGCTCCATGACCTCCATCCTCCTC 200
Db 33 -----LysTyrAlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeu 49
QY 201 GCCTACGATATCGGGGTGATGAGCGGGCGTGGCTGTACATCAAGAAGGACTTCAACATC 260
Db 50 GlyTyrAspIleGlyValMetSerGlyAlaValIleTyrIleLysLysAspLeuLysVal 69
QY 261 AGTGACGGGAAGTGGAGGTCTCATGGGCATATGAACTTCTACTCGCTCATCGGCTCC 320
Db 70 SerAspValGluIleGluValValGlyIleLeuAsnLeuTyrSerLeuIleGlySer 89
QY 321 TTCGGCGGGCGGACCTCGGACTGTGATCGCGCGGCGGTACACCATCGTGTTCGCGGCC 380
Db 90 AlalaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrIleValLeuAlaGly 109
QY 381 GTCATATTCCTCGCGGGGGSTTCCTCATGCGGGTTCGCGGTCAACTACGGCATGCTCATG 440
Db 110 AlaIlePhePheAlaGlyAlaLeuLeuMetGlyPheAlaProAsnTyrAlaPheLeuMet 129
QY 441 TTCGGCGGCTTCGTCGGCGGATCGGCTGGGCTACCGGCTCATGATCGCGCGGCTGATC 500

Db 130 PheGlyArgPheValAlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyr 149
QY 501 ACCGCGAGGTCTCGCGCGTCCGGTGGCTTCTGACGTCTTCCCGAGGTGTTCC 560
Db 150 ThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSerPheProGluValPhe 169
QY 561 ATCAACTTCGGCATCTCTCGCGTACGTCGAACTATGCTTTCTCCGTTGCCGCTG 620
Db 170 IleAsnAlaGlyIleLeuPheGlyTyrValSerAsnTyrGlyPheSerLysLeuProThr 189
QY 621 AACCTCGGTGGCATCATCTCGGATCGCGCGGCGCGTCCGCTCTGCTCGCGTCC 680
Db 190 HisLeuGlyTrpArgLeuMetLeuGlyValGlyAlaIleProSerIlePheLeuAlaIle 209
QY 681 ATGTGCTCGGCATCGCGGAGTCGCGGGTGGTGTGTCATGAAGGAGCGCTCGCGAC 740
Db 210 GlyValLeuAlaMetProGluSerProArgTrpLeuValMetGlnGlyArgLeuGlyAsp 229
QY 741 GCAAGGTGGTCTCTGGAGAAGACCTCCGACACGCGGAGGAGCGCGGAGGCTGCGC 800
Db 230 AlaArgLysValLeuAspLysThrSerAspSerLeuGluGluSerLysLysLeuArgLeuGly 249
QY 801 GACATCAAGGCGCGCGCGCATCTCTGAGGAGCTCGACGCGCGCTGCTGCTGCTGCTCC 860
Db 250 GluIleLysGluAlaAlaGlyIleProGluHisCysAsnAspAspIleValIleValLys 269
QY 861 AAGAGAGGAGCGGAAACGAGAGCGGTGTGGAAGAGGCTCATCTCTCCCGACCCCG 920
Db 270 LysArgSerGlnGlyGlnGlu-----ValTrpLysGlnLeuLeuLeuArgProThrPro 287
QY 921 GCCATGCGCGCATCTCTGCTGCGGATCGCATCCACTTCTTCCAGCATCGGTGGGC 980
Db 288 AlaValArgHisIleLeuMetCysAlaValGlyLeuHisPhePheGlnGlnAlaSerGly 307
QY 981 ATTCACTCGCTCTCTTACAGCCCTCTCGTGTTCAGAGCCCGCGGATTAAAGAACGAC 1040
Db 308 IleAspAlaValLeuTyrSerProArgIlePheGluLysAlaGlyIleThrAsnPro 327
QY 1041 AAACACTTCTGGGCACCATCTGGCGGTTCGCTGTCACCAAGAGCTTTTCATCTGTG 1100
Db 328 AspHisValLeuLeuCysThrValAlaValGlyPheValLysThrValPheIleLeuVal 347
QY 1101 GCGACTTCTTTCATCGACGGGTCTCGGGCGCGCGCTGTGCTGGGAGCAGCAGCGCGG 1160
Db 348 AlaThrPheMetLeuAspArgIleGlyArgArgProLeuLeuLeuThrSerValAlaGly 367
QY 1161 ATATCTCTCTCCCTCATCGGCTCGCGCGCGGCTCACGTCGTCGCGCAGACCCCGAC 1220
Db 368 MetValPheThrLeuAlaCysLeuGlyLeuGlyLeuThrIleIle---AspHisSerGly 386
QY 1221 GCCAAGATACCTTGGGCCATCGGCTAAGCATCGCTCCACCTCGCTCGCTCGCTTC 1280
Db 387 GluLysIleMetTrpAlaIleAlaLeuSerLeuThrMetValLeuAlaTyrValAlaPhe 406
QY 1281 TTCTTCATCGCTTGGCCCATCATCGTGGGTGTACAGCTCGAGATCTTCCGCTCCAG 1340
Db 407 PheSerIleGlyMetGlyProIleThrTrpValTyrSerSerGluIlePheProLeuGln 426
QY 1341 GTGCGCGGCTGGGTGCTCGCTCGGCTCGCGCAACCGCTCACCGCGGCGGTGATC 1400
Db 427 LeuArgAlaGlnGlyCysSerIleGlyValAlaValAsnArgValSerGlyValLeu 446
QY 1401 TCCATGACCTTCTCTGCTGTCTCAAGCCCATCACCATCGGCGGCGGTCTTCTCTCTAC 1460
Db 447 SerMetThrPheIleSerLeuTyrLysAlaIleThrIleGlyGlyAlaPhePheLeuPhe 466
QY 1461 TCGGCATCGCGCGCTCGCTGGGTGTCTTACACTTACCTACCTCCCGAGACCGCGGC 1520
Db 467 AlaAlaIleAlaValGlyTrpThrPhePhePheThrMetLeuProGluThrGlnGly 486
QY 1521 CGGACGCTGGAGGAGATGACCAAGCTTTCGGCGACACGCGCGCGCTCGGAATCAGAC 1580
Db 487 ArgThrLeuGluAspMetGluValLeuPheGlyLysPheTyrArgTrpArgLysAlaAsn 506

```

QY 1581 GAGCCAGCCAGGAGAGAGAGGTTGGAATGCGCCCGCCACTCACTGATCAAACTAACCG 1640
Db 507 AlaleuLeuLysGlnLysLysGlnValAspHisGlyAspGlyAsn-----AsnAsnPro 524

QY 1641 CAAAATCACCATAATC 1655
Db 525 AsnAsnProGlnIle 529

RESULT 4
Q9LS92 PRELIMINARY; PRT; 539 AA.
AC Q9LS92;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sugar transporter protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
EX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AB026854; BAB01812.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; F:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transport.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 539 AA; 58102 MW; 36B389284E5563A0 CRC64;

Alignment Scores:
Pred. No.: 5,16e-87 Length: 539
Score: 1625.50 Matches: 328
Percent Similarity: 75.76% Conservative: 69
Best Local Similarity: 62.60% Mismatches: 116
Query Match: 44.98% Indels: 11
DB: 10 Gaps: 6

US-10-051-902a-21 (1-2017) x Q9LS92 (1-539)

QY 60 CCAGAGGAGCGCTCTCTCTCTGACACCGAGATGCTTCCCGCGCTGCCGGAG 119
Db 6 ProGluLeuThrAlaProSerProProPro-----ValLysHisValProGlu 22

QY 120 GCCCTCCGCGGAGAGAG-----AAGGCAACGTCGGTTCGGCTTCGGCTCGGCATC 173
Db 23 SerValLeuProAlaLysProProLysArgAsn---AsnTyrAlaPheAlaCysAlaIle 41

```

```

174 CTCGCTCCATGACCTCCATCTCTCTCGGTACGATATCGGGGTGATAGCGGGCGGTGC 233
Db 42 LeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaMet 61

234 CTGTACATCAAGAGAGACTTCAACATCAGTCAGCGGAAGGTGAGGTTCTCATGGCATA 293
Db 62 IleTyrIleLysArgAspLeuLysIleAsnAspLeuGlnIleGlyIleLeuAlaGlySer 81

294 CTGAACCTCTACTCGCTCATCGGCTCTCGGGGGGGCGGACGTCGAGCTGGAATCGGC 353
Db 82 LeuAsnIleTyrSerLeuIleGlySerCysAlaAlaGlyArgThrSerAspIleGly 101

354 CGCGGTACACCATCGTGTTCGCCCGCTCATATTCTTCGGGGGGSGTTCCTCATGGG 413
Db 102 ArgArgTyrThrIleValLeuAlaGlyAlaIlePhePheAlaGlyAlaIleLeuMetGly 121

414 TTCGGCGTCACTACGACCTCATGTTCGGCCCTTCGTGGCCGGCATCGCGTGGGC 473
Db 122 LeuSerProAsnTyrAlaPheLeuMetPheGlyArgPheIleAlaGlyIleGlyValGly 141

474 TACGCGCTCATCATCGCGCGGTGTACACCGCGGAGGTGTCCCGCGCTCGCGCGGTGC 533
Db 142 TyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGly 161

534 TTCCTGAGCTGTTCCCGGAGGTGTTCATCACTTCGGCATCTCTCTCGGTACGTCGT 593
Db 162 PheLeuAsnSerPheProGluValPheIleAsnAlaGlyIleMetLeuGlyTyrValSer 181

594 AACTATGTTCTCCCGCTTCGCGTGAACCTCGGGTGGCGCATCATCTCGGCATCGC 653
Db 182 AsnLeuAlaPheSerAsnLeuProLeuLysValGlyTyrArgLeuMetLeuGlyIleGly 201

654 GCGGCGCTCGCTGCTCTCGCTCATGTGTCTCGCATCGCGAGTCGCGCGGTGG 713
Db 202 AlaValProSerValIleLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrp 221

714 CTGGTCATGAAGGAGCGCTCGCGACCGCAAGTGTGTCTGGAGAAGACCTCCACACG 773
Db 222 LeuValMetGlnGlyArgLeuGlyAspAlaLysArgValLeuAspLysThrSerAspSer 241

774 GCGGAGGAGCGCGCGGAGCGCTCGCGCACATCAAGCGCGCGCGCGCATCCCTCAGGAG 833
Db 242 ProThrGluAlaThrLeuArgLeuGluAspIleLysHisAlaAlaGlyIleProAlaAsp 261

834 CTCGACGCGGAGCTGTGTACCGTCCCAAGAGAGGAGCGGAAACGAGAAGCGGTGTGG 893
Db 262 CysHisAspAspValValGlnValSerArgArgAsnSerHisGlyGlu---GlyValTrp 280

894 AAGGAGCTCATCTCTCCCGACCCCGCGCATCGCGGCATCTCTCTCGCGGATCGGC 953
Db 281 ArgGluLeuLeuIleArgProThrProAlaValArgValMetIleAlaIleGly 300

954 ATCCACTTCTCCAGCATCGGTGGCGATTCACTCCGTCGTCTTCTACAGCCCTCTCGTG 1013
Db 301 IleHisPhePheGlnAlaSerGlyIleAspAlaValValLeuPheSerProArgIle 320

1014 TTCAGAGCCCGGATTAAACGACGACAAACACTTCTTGGGCACCACTTGGCCGTCCTCGT 1073
Db 321 PheLysThrAlaGlyLeuLysThrAspHisGlnGlnLeuLeuAlaThrValAlaValGly 340

1074 GTCACCAAGAGGCTTTTCATCTTGTGGCGACTTCTTCTCATCGCGGTGGCGGGCGG 1133
Db 341 ValValLysThrSerPheIleLeuValAlaThrPheLeuLeuAspArgIleGlyArgArg 360

1134 CGCTGTTCGTGGGAGCAGCGGGGGGATATCTCTCTCTCATCGGCTCGCGCCCGG 1193
Db 361 ProLeuLeuLeuThrSerValGlyGlyMetValLeuSerLeuAlaLeuAlaGlyThrSer 380

1194 CTCACCGCTCGTCGCGCAGCACCCCGACCGCAAGTACCTTGGCCATCGGCTAAGCATC 1253
Db 381 LeuThrIleAspGln---SerGluLysLysValMetTrpAlaValAlaValAlaIle 399

```


QY 1359 TCCTCGCGCTCGCGCCACCGGTCCACCGCGGTCTCATCTCCATGACCTTCCTGTG 1418
 Db 429 ALAAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448
 QY 1419 CTGTCGAAGCCATCATCTCGCGCGAGCTTCTCTCTACTCCGCGCATCGCGCGCTC 1478
 Db 449 LeuThrArgAlaIleThrIleGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal 468
 QY 1479 GCCTGGGTCTTCTACACTACTCCCGGAGACCGCGCGCGAGCTGGAGGAGATG 1538
 Db 469 GlyTrpIlePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488
 QY 1539 AGCAAGCTGTTCGCGACACCGCGCGCTCGGAATCAGACGACCGCAGGAGAAAG 1598
 Db 489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla-----Ser 502
 QY 1599 AAGAAGTGAAGTGGCGGCACCTAACTGATCAAACTAACCGCAAAATCACCAATCCTA 1658
 Db 503 LysAlaValGlu-----AsnGluAsnGlyGlnValAla 513
 QY 1659 AGGGTTTCTTGC AAAACCGTGTG 1682
 Db 514 GlnValGlnLeuGlyThrAsnVal 521

RESULT 6
 Q84KI7 PRELIMINARY; PRT; 509 AA.
 AC Q84KI7;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Sorbitol transporter.
 OS Prunus cerasus.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids;
 OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=140311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fruit;
 RX MEDLINE=22578918; PubMed=12692316;
 RA Gao Z., Maurousset L., Lemoine R., Yoo S.-D., van Nocker S.,
 RA Loescher W.;
 RT "Cloning, Expression, and Characterization of Sorbitol Transporters
 RT from Developing Sour Cherry Fruit and Leaf Sink Tissues.";
 RL Plant Physiol. 131:1566-1575(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fruit;
 RA Zhifang G., Loescher W.H., Lemoine R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF482011; AA039257.1;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub.transporter.
 DR InterPro; IPR003663; Sugar transpt.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; sugar trf_1.
 DR PRINTS; PA00171; SUGRTNSPOT.
 DR Trifam; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 SQ SEQUENCE 509 AA; 55407 MW; EA25ED065CE884C8 CRC64;

Alignment Scores:
 Pred. No.: 6,17e-86 Length: 509
 Score: 1607.00 Matches: 319
 Percent Similarity: 78.48% Conservative: 53

Best Local Similarity: 67.30% Mismatches: 98
 Query Match: 44.47% Indels: 4
 DB: 10 Gaps: 3
 US-10-051-902A-21 (1-2017) x Q84KI7 (1-509)
 QY 132 AAGAAGAGGCAACGTCGCGGTTCGCTTCGCTCGGCCATCTCGCTCCATGACCTCC 191
 Db 20 LysProLysArgAsnLeu---TyrAlaIleGlyCysAlaIleLeuAlaSerMetThrSer 38
 QY 192 ATCTCTCTCGCTACGATATCGGGTGATGACCGGGGCGTGCCTGTACATCAAGAGGAC 251
 Db 39 IleLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaSerIleTyrIleGlnLysasp 58
 QY 252 TTCACATCAGTCAGCGGAGGTTCATGGGCATCTGAACTCTACTTCGCTC 311
 Db 59 LeuLysIleSerAspValGluValGluIleLeuIleGlyIleLeuAsnLeuTyrSerLeu 78
 QY 312 ATCGCTCTCTCGCGCGCGGAGACGTGCGACTCGGATCGCGCGCGGTACACATCTGTG 371
 Db 79 IleGlySerAlaAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrIleVal 98
 QY 372 TTCGCGCGCGTCATATTCTTCGCGGGGGGTTCCTCATGGGGTTCGCGCTCAACTACGCC 431
 Db 99 PheAlaGlyAlaIlePhePheThrGlyAlaLeuLeuMetGlyLeuAlaThrAsnTyrAla 118
 QY 432 ATGCTCATGTTTCGCGCGCTTCGTGCGCGCATCGCGTGGGTACGCGCTCATGATCGCG 491
 Db 119 PheLeuMetValGlyArgPheValAlaGlyIleGlyValGlyTyrAlaLeuMetIleAla 138
 QY 492 CGGTGTACACCGCGAGGTGTCGCGCGGTGTCGCGCGGTCTCTCCTCAGCTCGTCCCG 551
 Db 139 ProValTyrAsnAlaGluValSerProAlaSerSerArgGlyAlaLeuThrSerPhePro 158
 QY 552 GAGGTGTTTCATCAACTTCGCGCATCTCGGTACGTCTCGAACTATGCTTTCCCGCG 611
 Db 159 GluValPheValAsnIleGlyIleLeuLeuGlyTyrValAlaAsnTyrAlaPheSerGly 178
 QY 612 TTGCGCTGAACCTCGGTGGCGATCATGTCTGCGCATCGCGCGCGCGCTCGCTCGTG 671
 Db 179 LeuProIleAspLeuGlyTyrArgLeuMetLeuGlyValGlyValPheProSerValIle 198
 QY 672 CTCGCGCTCATGCTCGGATCGCGGTGCGCGGTGCTGCTCATGAAGGAGCGC 731
 Db 199 LeuAlaValGlyValLeuSerMetProGluSerProArgTrpLeuValMetGlnGlyArg 218
 QY 732 CTCGCGGACCGCAAGTGTGCTGGAGAAGACCTCCGACACGCGGAGAGCGCGGAG 791
 Db 219 LeuGlyGluAlaLysGlnValLeuAspLysThrSerAspSerLeuGluGluAlaGlnLeu 238
 QY 792 CGCTGGCGGACATCAAGCGCGCGCGCATCTCGAGAGCTCGACGGCGACGTGTG 851
 Db 239 ArgLeuAlaAspIleGlyGluAlaAlaGlyIleProGluHisCysValGluAspValVal 258
 QY 852 ACCGTCCCCAAGAGAGGAGCGGAAACGAGAGCGGTGTGGAAGGAGCTCATCTGTCC 911
 Db 259 GlnValProLysHisSerHisGlyGluGlu-----ValTrpLysGluLeuLeuHis 276
 QY 912 CGACCGCGGCATCGCGCGCATCTCTGTGTCGCGGATCGGCATCTCTTCAGCAT 971
 Db 277 ProThrProValArgHisIleLeuIleAlaAlaIleGlyPheHisPhePheGlnGln 296
 QY 972 GGTGGGCACTTCACTCGTGTCTTCTACAGCCCTCTCTGTTTCAAGAGCGCGCGATTA 1031
 Db 297 LeuSerGlyIleAspAlaLeuValLeuTyrSerProArgIlePheGluLysAlaGlyIle 316
 QY 1032 ACGAAGCAAAACACTTCTTGGGACCACTTGGCGGTTCGGGTGTCCCAAGAGGCTTTTC 1091
 Db 317 ThrAspSerThrLeuLeuLeuAlaThrValAlaValGlyPheSerLysThrIlePhe 336
 QY 1092 ATCTGTGTGGCGACTTCTTTCATCGACGCGCTCGCGCGCGCGCTGTGTGGCGGACG 1151
 Db 337 ThrLeuValAlaIleGlyPheLeuAspArgValGlyArgProLeuLeuLeuThrSer 356


```
Db 267 ILfYLYsValGluLysThrLysLleLysSerGlyAsnAlaValTrpLysGluLeuPhe 286
QY 906 CTCTCCCGACCCCGCCATCGCGGCATCTCTCTCGGGATCGGCATCGGCATCTCTTC 965
Db 287 PheAsnProThrProAlaValArgAlaValAlaValAlaGlyLleGlyLleHisPhe 306
QY 966 CAGCATCGGTGGGATTCATCTCCGCTCTCTACAGCCCTCTCGTGTGAAGAGCCCC 1025
Db 307 GlnGlnAlaSerGlyLleAspAlaValValLeuTrpSerProArgLysPheGlnSerAla 326
QY 1026 GGATTAAACGAACACAAACACTCTCTGGGCACCATCTGGCGCTCGCTCACCAGAGG 1085
Db 327 GlyLleThrAsnAlaArgLysGlnLeuAlaThrValAlaValGlyValLysThr 346
QY 1086 CTCTTCATCTGTGGCGACTTCTCTCATCGAGCGCTGGGGGGCGCGCTGTCTGCTG 1145
Db 347 LeuPheLleLeuValAlaThrPheGlnLeuAspLysTrpGlyArgArgProLeuLeuLeu 366
QY 1146 GGCAGACGGCGGGATATCTCTCTCTCATCGGCTCGGCGCGGCTCACCGTCGTC 1205
Db 367 ThrSerValGlyLleMetLleAlaLleLeuThrLeuAlaMetSerLeuThrValLle 386
QY 1206 GGCAGACACCCGACCCAGATACCTTGGGCGCATCGGCTAAGCATCGCTCCACCTC 1265
Db 387 ---AspHisSerHisLysLleThrTrpAlaLleAlaLeuLysLleThrMetValCys 405
QY 1266 GCCTAGCTGCTCTCTCTCCATCGCGCTGGCGCCCATCACGCGGTGTACAGCTCGGAG 1325
Db 406 AlaValValAlaSerPheSerLleGlyLeuGlyProLleThrTrpValTrpSerSerGlu 425
QY 1326 ATCTTCCCGCTCCAGGTGCGCGGCTGGGTGCTGCTCGGCTCGCGCCGCCAACCGCTC 1385
Db 426 ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 445
QY 1386 ACAGCGGCTCATCTCCATCATCTCTCTCTCTCGCTCCAGGCCCATCACCTCGCGGC 1445
Db 446 ValSerGlyValLleSerLlePhePheLeuProLeuSerHisLysLleThrThrGlyGly 465
QY 1446 AGTTCTCTCTCTACTCCGCGATCGCGCGCTCGCTCGGTGTCTTCTACACCTACCTC 1505
Db 466 AlaPhePheLeuPheGlyLleAlaLleAlaLleAlaTrpPhePheLeuThrPheLeu 485
QY 1506 CCGGAGACCCCGCGCGGACGCTGGAGAGATGAGCAAGTGTTCGGCGGAC----- 1556
Db 486 ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTrp 505
QY 1557 -----ACGGCGCGCGCTCGGAATCATGAGCGAGCGCGC----- 1589
Db 506 ArgGluSerPheProGlyAsnLysSerAsnAsnAspGluAsnSerThrArgLysGlnSer 525
QY 1590 -----AAGGAGAGAGAGAGGTGGAAATGGCGGCCACT 1622
Db 526 AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThr 538

RESULT 10
Q9XIH7 PRELIMINARY; PRT; 511 AA.
AC Q9XIH7
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative sugar transporter.
GN AT2G16120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_taxid=3702;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
```

```
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vankken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RL thaliana";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AC007134; AAD26954.1; -.
DR PIR; H84536; H84536.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005311; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MPS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS50850; MPS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
DR Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 511 AA; 54758 MW; E646A59316C6A54B CRC64;

Alignment Scores:
Pred. No.: 2,688-84 Length: 511
Score: 1579.00 Matches: 306
Percent Similarity: 75.10% Conservative: 77
Best Local Similarity: 60.00% Mismatches: 121
Query Match: 43.69% Indels: 6
DB: 10 Gaps: 2

US-10-051-902A-21 (1-2017) x Q9XIH7 (1-511)
QY 96 ATGGTTCGCGCGCTCGCGAGCGCGCTCGC-----CCGAGAGAGAG 140
Db 1 MetAsnSerSerGlyValGluGlnGlyValLleAlaGluSerGluProArgGly 20
QY 141 GGCACAGTCGGTTCGCTTCGCTCGCCATCTCGCTCCATCCATCCATCCATCCCTC 200
Db 21 AsnArgSerArgTrpAlaPheAlaCysAlaLleLeuAlaSerMetThrSerLleLeu 40
QY 201 GGCTACGATATCGGGTGTATGAGCGGGCGTGTATCATCAACAAGAGAGCTTCAACATC 260
Db 41 GlyTyraAspLleGlyValMetSerGlyAlaSerLlePheLleLysAspAspLeuLysLeu 60
QY 261 ACTGACGGAGAGTGGAGGTTCATGGGCGTACTGGAACCTCTACTCGCTCATCGCTCC 320
Db 61 SerAspValGlnLeuLleLeuMetGlyLleLeuAsnLleTyrSerLeuValGlySer 80
QY 321 TTCGCGGGGGGCGAGCTCGGACTGGATCGCGCGGCTACACCATCGTGTTCGCCGCC 380
Db 81 GlyAlaAlaGlyArgThrSerAspTrpLeuGlyArgArgTrpThrLleValLeuAlaGly 100
QY 381 GTCATATTTCTTCGCGGGGGSGTTCCTCATAGGGGTTCGCGGTCAACTAGCCATGCTCATG 440
Db 101 AlaPhePhePheCysGlyAlaLeuLeuMetGlyPheAlaThrAsnTyrProPheLleMet 120
QY 441 TTCGCGCGGCTTCGCGCGGCGATCGCGCTGCGCTACGCGCTACGCGCTCGCGGTGTAC 500
Db 121 ValGlyArgPheValAlaGlyLleGlyValGlyTyrAlaMetMetLleAlaProValTrp 140
```



```
Db 20 GlyAsnSerArgPheAlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleIle 39
QY 198 CTCGGTACATATCGGGGTGATGACGGGGGCTGCTGTACATCAAGAGACTTCAAC 257
Db 40 LeuGlyThrAspIleGlyValMetSerGlyAlaAlaIlePheIleIleAspLeuLys 59
QY 258 ATCAGTACGGGAGGTGAGGTTCATCAGGCATACCTCAAGCTCTACCTCGCTACCGC 317
Db 60 LeuSerAspValGlnLeuGlnIleLeuMetGlyIleLeuAsnIleIleIleIleGly 79
QY 318 TCCTTCGGCGGGGGGAGCTCGACTGATCGCGCGGGGTACACCATCGTGTTCGCC 377
Db 80 SerGlyAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTrpThrIleValLeuAla 99
QY 378 GCCGTATATCTTCGGGGGGGGTTCCTCATGGGGTTCGGCTCACTACCCATGCTC 437
Db 100 GlyPhePhePheCysGlyAlaLeuLeuMetGlyPheAlaThrAsnTrpProPheIle 119
QY 438 ATCTTCGGCGGCTTCGTGGCGGCATCGCGCTCGGCTCATGCTCATGCTCGCGGTG 497
Db 120 MetValGlyArgPheValAlaGlyIleGlyValGlyTrpAlaMetMetIleAlaProVal 139
QY 498 TACACCGCGAGGTGTCGGCGGTTCGGCGGTTCCTGACGTCGTTCGGGAGGTG 557
Db 140 TyrThrThrGluValAlaProAlaSerSerArgGlyPheLeuSerSerPheProGluIle 159
QY 558 TTCATCACTTCGGCATCTCGTCCGGGTAGCTCGAATGATCTTCTCCGCTTCGCCG 617
Db 160 PheIleAsnIleGlyIleLeuGlyTrpValSerAsnTrpPhePheAlaIleLeuPro 179
QY 618 CTGAACCTCGGGTGGCGCATCATGCTCGGCATCGCGCGCGCGCTCGCTCGCTCGCG 677
Db 180 GluHisIleGlyTrpArgPheMetLeuGlyIleGlyAlaValProSerValPheLeuAla 199
QY 678 CTATGTCGTTCGGCATCGCGGTTCGGCGGTTCGTGTCATGAGGAGCGCTCGCG 737
Db 200 IleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGlnGlyArgLeuGly 219
QY 738 GAGCGCAAGGTGCTGTGGAGAGACCTCCGACACGGCGGAGGAGCGCGCGCGCTCG 797
Db 220 AspAlaPheLysValLeuAspLysThrSerAsnThrLysGluAlaIleSerArgLeu 239
QY 798 GCCGACATCAAGCGCGCGCGCATCTCTGAGAGCTCGACGGCGAGCTGTGACCGTC 857
Db 240 AsnAspIleLysArgAlaValGlyIleProAspAspMetThrAspValIleVal 259
QY 858 CCCAAGAGGAGCGGAGAGACGAGCGGTGTGGAAGAGCTCATCTGCTCCCGGAC 917
Db 260 ProAsnLysLysSerAlaGly---LysGlyValTrpLysAspLeuValArgProThr 278
QY 918 CCGGCCATCGCGGCATCTGCTGTCCGGATCGGCATCCACTTCTTCCAGCATCGGTG 977
Db 279 ProSerValArgHisIleLeuIleAlaCysLeuGlyIleHisPheSerGlnGlnAlaSer 298
QY 978 GGATTCATCTCTCTCTACAGCCCTCTGCTTCAAGAGCCCCGGATTAAAGAAC 1037
Db 299 GlyIleAspAlaValValLeuTrpSerProThrIlePheSerArgAlaGlyLeuLysSer 318
QY 1038 GACAAACATCTTTCGGGCACCATCTGCGCTGTCACCAAGAGCTTTTCATCTTG 1097
Db 319 LysAsnAspGlnLeuLeuAlaThrValAlaValGlyValValLysThrLeuPheIleVal 338
QY 1098 TTGGCGATTTCTTCATCAGCGGTTCGGCGGGCGCGCTGTGTCTGGGCACACGCGG 1157
Db 339 ValGlyThrCysLeuValAspArgPheGlyArgArgAlaLeuLeuLeuThrSerMetGly 358
QY 1158 GGGATATCTCTCTCTCATCGGCTCGCGCGCGGTCACTCGCTCGCGCGCGCGCGG 1217
Db 359 GlyMetPhePheSerLeuThrAlaLeuGlyThrSerLeuThrValIleAspArgAsnPro 378
QY 1218 GACGCCAAGATACCTTGGCGCATCGGCTAAGCATCGCTCCACCTCGCTCGCTCGCG
```

```
379 GlyGlnThrLeuLysTrpAlaIleGlyLeuAlaValThrValMetThrPheValAla 398
QY 1278 TTCCTTCATCCATCGCTTGGCCCATCAGTGGGTGTACAGTCGAGATCTTCCCGTC 1337
Db 399 ThrPheSerLeuGlyAlaGlyProValThrTrpValTrpAlaSerGluIlePheProVal 418
QY 1338 CAGTGTGGCGCGCTGCTGCTCGCTCGCGCTCCGCCCAACCGCTACACCGCGGTC 1397
Db 419 ArgLeuArgAlaGlnGlyAlaSerLeuGlyValMetLeuAsnArgLeuMetSerGlyIle 438
QY 1398 ATCTTCATGACTTCTCTGCTGCTGCTCAAGCCATCAGCTCGGCGGAGCTTCTCTC 1457
Db 439 IleGlyMetThrPheLeuSerLeuSerLysGlyLeuThrIleGlyAlaPheLeuLeu 458
QY 1458 TACTCCGGCATCGCGCGCTCGCTCGCTGCTGTCTTCTACACTACCTCCCGGAGCCGC 1517
Db 459 PheAlaGlyValAlaValAlaAlaTrpValPhePheThrPheLeuProGluThrArg 478
QY 1518 GCGCGGAGCTGTGAGGAGATGAGCAAGCTGTTCGGCGACACGCGCGCTCGGATCA 1577
Db 479 GlyValProLeuGluGluIleGluSerLeuPheGlySerTyrSerAlaAsnLysLysAsn 498
QY 1578 GACGAGCCAGCCAGGAGAGAGAGAGAGAGTGGAA 1610
Db 499 AsnValMetSerLysGlyLysGlnValValAsp 509

RESULT 12
Q84RI2 PRELIMINARY; PRT; 491 AA.
AC Q84RI2, 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Sorbitol transporter.
GS SOTI.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=cv. Mutsu; TISSUE=Fruit;
RA ZhiFang G., Jayanty S., Beaudry R.M., Loescher W.H.;
RT "Cloning and characterization of apple fruit sorbitol transporter.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237400; AAC8964.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub. transporter.
DR InterPro; IPR003663; Sugar transp.
DR InterPro; IPR005829; Sug. transporter.
DR Pfam; PF00083; sugar tr. 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 491 AA; 53310 MW; 3F9AE2E85CCB040A CRC64;

Alignment Scores:
Pred. No.: 8,578-81 Length: 491
Score: 1519.00 Matches: 288
Percent Similarity: 77.68% Conservative: 81
Best Local Similarity: 60.63% Mismatches: 104
Query Match: 42.03% Indels: 2
DB: 10 Gaps: 1

US-10-051-902A-21 (1-2017) x Q84RI2 (1-491)
```



```
QY 162 GCCTGCGCATCTCGCTCCATGACCTCCATCCTCTCGGCTACGATATCGGGGTGATG 221
Db 38 GlyCysAlaIleLeuAlaSerMetThrSerValLeuLeuGlyTyrAspIleGlyValMet 57
QY 222 AGCGGGCGTCCGTGATCAACAAGACTTCAACATCAGTACGGGAGGTGGAGTT 281
Db 58 SerGlyAlaAlaIleTyrIleLysAspGlnLeuHisValSerAspValLysLeuGluIle 77
QY 282 CTCATGGGCATACGAACCTCTACTCGCTCATCGCTCCCTTCGCGGGGGCGGAGCGTCG 341
Db 78 ValValGlyIleIleAsnPheSerLeuValGlySerAlaLeuAlaGlyAraGlyThrSer 97
QY 342 GACTGGATCGCGCGCGGTACACCATCGTGTTCGCCCGCTCATATCTTCGCGGGGGSG 401
Db 98 AspTrpIleGlyArgArgTyrThrMetValLeuAlaGlyAlaIlePhePheValGlyAla 117
QY 402 TTCCTCATGGGTTCGCGGTCAACTAGCCATGCTCATGTCGCGCGCTTCGTCGGCGGC 461
Db 118 IleLeuMetGlyPheAlaThrAsnTyrSerPheLeuMetPheGlyArgPheValAlaGly 137
QY 462 ATCGCGGTGGGTACGCGCTCATGATCGCGCGGTGTACACCGCGAGGTGTGCGCGGC 520
Db 138 IleArgSerArgLeuCysProHisAspCysSerGlyValHisGlyAraGlyPhePheSer 157
QY 521 GTCGCGCGTGGTTCCTCGTACGTGTCCTCGGAGGTGTTTCATCAACTTCGCGCATCTGCT 580
Db 158 IleIleSerTrpIleSerTyrPheIleProGlySerPheIleAsnIleGlyValLeuLe 177
QY 581 CGGTACGCTCGAATATGTTTCTCCGCTTCCGCTGACCTCGGTCGGGTGGCGCATCAT 640
Db 177 uGlyTyrValSerAsnTyrAlaPheSerLysLeuProAlaAsnLeuGlyTrpArgPheMe 197
QY 641 GCTCGGATCGCGCGCGGTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 700
Db 197 tLeuGlyIleGlyAlaIleProSerIleGlyLeuAlaIleGlyValLeuGlyMetProG 217
QY 701 GTCCCGCGGTGGTGTGTATGAAGGACGCTCGCGGACGCCCAAGTGTGTGAGAA 760
Db 217 uSerProArgTrpLeuValMetLysGlyArgLeuGlyGluAlaArgGlnValLeuAsp 237
QY 761 GACTCCGACACGCGGAGGAGCGCGGAGCGCTGCGGACATCAAGCGCGCGCGCGG 820
Db 237 sThrSerAspSerLysGluLeuArgLeuArgLeuSerAspIleLysGlnAlaAlaG 257
QY 821 CATCCCTGAGAGCTCAGCGGACGTGTGACCGTCCCGACAGAGGAGGCGGAAACGA 880
Db 257 yIleProGluGluCysAsnAspIleValValMetProLysArg-----ArgAsnAs 275
QY 881 GAAGCGGTGTGAAGAGAGTCTCTCTGTCCTCCCGACCCCGCATGCGCGCATCTCTGCT 940
Db 275 pGluAlaValTrpLysGluLeuLeuHisProThrProSerValArgHisAlaPheI 295
QY 941 GTCCGGATCGCATCCACTTCTTCAGCATGCGTGGGATTCATCTCGCTGCTCTCTTA 1000
Db 295 eAlaGlyValGlyLeuHisPhePheGlnGlnSerSerGlyIleAspAlaGlyLeuTy 315
QY 1001 CAGCCCTCTGCTGTTCAAGAGCCCGGATTAAACGAACGACAAACACTTCTTGGGCACC 1060
Db 315 rSerProArgIlePheGluLysAlaGlyIleThrSerThrAspLeuLysLeuAlaI 335
QY 1061 TTGCGCGTTCGTTCCACCAAGAGGCTTTTCATCTTCTTGGGACTTCTTCATCGACGG 1120
Db 335 rIleAlaValGlyIleSerLysThrLeuPheIleLeuValAlaThrPheLeuLeuAsp 355
QY 1121 CGTCGGCGGCGCGGTGTGTGTCGACGACGCGGCGGATATCTCTCTCCCTCATCGG 1180
Db 355 gIleGlyArgArgProLeuLeuLeuThrSerMetGlyGlyMetIleIleSerLeuThr 375
QY 1181 CCTCGCGCGCGGTCTACCGTCTGTCGCG-----1208
Db 375 uLeuGlyThrSerLeuAlaValIleAlaThrGlnIleThrGlnPheIleGlyProGlyHi 395
QY 1209 ----CAGACCCCGACGCCAAGATACCTTGGGCCATCGGCTAAGCATCCGCTCCACCT 1264
```

Search completed: June 30, 2004, 18:52:02
Job time : 154.5 secs

```
Db 395 sTrpGlnSerSerGluValTyrHisThrTrpAlaVal-----407
QY 1265 CCCTTACGTGCGCTTCTTCTCCATCGGCTTGGCCCATCACGTGGGTCTACAGCTCGGA 1324
Db 408 -----PhePhe-----TrpLeuGlyProIleAlaIleTrpGlyTyrSerSerGI 421
QY 1325 GATCTTCCGCTCCAGGTGCGCGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1384
Db 421 uValPheProLeuArgLeuArgAlaGlnGlyCysSerIleGlyValAlaValAsnArgGI 441
QY 1385 CACCGCGCGCTCATCTCCATCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1444
Db 441 yThrSerGlyIleIleSerMetThrPheLeuSerLeuTyrLysAlaIleSerIleAlaGI 461
QY 1445 CAGCTTCTCTCTACTCTCGGCTCGGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1504
Db 461 yAlaPheTyrLeuPheAlaAlaIleAlaGlyValAlaTrpIlePheIlePheThrLeuLe 481
QY 1505 CCGGAGAGCCCGCGCGGACGCTGGAGAGATGAGCAAGCTGTTCGGC 1553
Db 481 uProGluThrGlnGlyArgSerLeuGluGluMetGlyLeuLeuPheGly 497
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 02:11:44 ; Search time 12597 Seconds
(without alignments)
6939.978 Million cell updates/sec

Title: US-10-051-902a-21

Perfect score: 2017

Sequence: 1 ctctactgaagctcgtgc.....aaaaaaaaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_btg.*

3: gb_in.*

4: gb_cm.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2016.6	100.0	2017	6	AR208575	Sequence
2	1887	93.6	1961	8	AK102676	Oryza sat
3	1880.6	93.2	1962	8	AK059423	Oryza sat
4	1783	88.4	25270	8	AP004185	Oryza sat
5	1660.6	82.3	1860	8	AF416867	Oryza sat
6	1120.2	55.3	2109	8	BT009264	Trifolium
7	1094.6	54.3	2089	6	AR208577	Sequence
8	1032.2	51.2	1620	8	AK119464	Oryza sat
9	1005	49.8	1914	6	AR208574	Sequence
10	988	49.0	1872	8	BT009301	Trifolium
11	985.8	48.9	25270	8	AP004185	Oryza sat
12	981.6	48.7	1872	6	AR208578	Sequence
13	909.2	45.1	193577	8	AP004365	Oryza sat
14	759.4	37.6	150337	8	AC116426	Genomic s
15	759.4	37.6	181163	8	AC115687	Genomic s
16	739.6	36.7	2300	8	AK070417	Oryza sat
17	738	36.6	2167	8	AK103915	Oryza sat
18	729.2	36.2	2252	8	AK099273	Oryza sat
19	687.8	34.1	139123	8	AC145480	Genomic s
20	657.4	32.6	1998	8	AK121279	Oryza sat
21	644.4	31.9	1927	8	AF482011	Prunus ce
22	628.6	31.2	2032	8	AY100638	Prunus ce
23	620	30.7	141050	8	AC093093	Oryza sat
24	620	30.7	147706	8	AC083945	Oryza sat
25	620	30.7	301337	8	AE017082	Oryza sat
26	603.2	29.9	1848	8	AY136668	Orobancha
27	600.4	29.8	2012	8	AY237400	Malus x d
28	590.2	29.3	1729	8	AK061613	Oryza sat
29	590.2	29.3	1857	8	AK099395	Oryza sat
30	588	29.2	1784	8	AY237401	Malus x d
31	576.2	28.6	1987	8	PM532589	Plantago
32	575.2	28.5	1853	6	AR208576	Sequence
33	573.6	28.4	1614	8	AK106745	Oryza sat
34	561.2	27.8	1894	8	GMA563367	Glycine m
35	548.8	27.2	1638	8	AK119762	Oryza sat
36	546.8	27.1	133331	2	OSJUN00017	Sequence
37	545.2	27.0	103264	2	OSJUN00049	Sequence
38	512.6	25.4	1946	8	AK121039	Oryza sat
39	510.8	25.3	1891	8	PM532590	Plantago
40	504.6	25.0	162766	8	AC125782	Oryza sat
41	504.6	25.0	171946	8	AC145349	Oryza sat
42	504.6	25.0	348970	2	AC138608	Oryza sat
43	477.2	23.7	2054	8	BUV64903	Beta vulgar
44	477.2	23.7	2056	6	AX353437	Sequence
45	469.2	23.3	1925	6	AX353438	Sequence

ALIGNMENTS

RESULT 1
AR208575
LOCUS AR208575 2017 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 21 from patent US 6383776.
ACCESSION AR208575
VERSION AR208575.1 GI:21509765
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2017)
AUTHORS Allen, S.M., Hitz, W.D., Kinney, A.J. and Tingey, S.V.
TITLE Plant sugar transport proteins
JOURNAL Patent: US 6383776-A 21 07-MAY-2002;
FEATURES Location/Qualifiers

[illegible]

609	Db	GGTGGCGCATATGCTCGGCATCGCGCGCGCGTCCGTGCTGCTCGCGTCAATGATGC	668
688	Qy	TCGGCATCGCGAGTCGCGCGGTGGCTGCTCATGAAGGACGCTCGCGACGCCAAGG	747
669	Db	TCGGCATCGCGAGTCGCGCGGTGGCTGCTCATGAAGGACGCTCGCGACGCCAAGG	728
748	Qy	TGTTGCTGGAGAAAGCTTCGACACGCGGAGGAGCGCGGAGCGCTGGCCGACATCA	807
729	Db	TGTTGCTGGAGAAAGCTTCGACACGCGGAGGAGCGCGGAGCGCTGGCCGACATCA	789
808	Qy	AGGCCGCGCGCGCATCCCTAGAGAGCTCGACGCGAGCTGTGTACCGTCCCAAGAGAG	867
789	Db	AGGCCGCGCGCGCATCCCGAGAGCTCGACGCGAGCTGTGTACCGTCCCAAGAGAG	848
868	Qy	GGAGCGAAACGAGAAGCGGTGTGGAAGAGCTCATCTGTCCCCGACCCCGGCCATGC	927
849	Db	GGAGCGAAACGAGAAGCGGTGTGGAAGAGCTCATCTGTCCCCGACCCCGGCCATGC	908
928	Qy	GGCGCATCTGCTGCTCGCGGATCGGCATCCACTTCTTCAGCATGCGTGTGGGCAATCACT	987
909	Db	GGCGCATCTGCTGCTCGCGGATCGGCATCCACTTCTTCAGCATGCGTGTGGGCAATCACT	968
988	Qy	CCGTGCTCTTCTACAGCCCTCTCGTGTTCAGAGCCCCGGATTAAACGAAAGCAACCACT	1047
969	Db	CCGTGCTGCTCTACAGCCCGCGCGTGTTCAGAGCGCGGCATCACCGACGACCAACACC	1028
1048	Qy	TCTTGGGCAACAATTGGCCGTTTGGTGTACCAAGAGGCTTTTCATCTTGTGTGGCACTT	1107
1029	Db	TCTTGGGCAACAATTGGCCGCTCGGTGTACCAAGAGCGCTTTCATCTCTGTGGCGACCT	1088
1108	Qy	TCTTTCAGAGGCTCGCGCGCGCGCTGTTGCTGGCAGCACGCGGCGGGATATCC	1167
1089	Db	TCTTCTCAGCGCTCGCGCGCGCGCTGCTGTAGCAGCACGCGGCGGGATATCC	1148
1168	Qy	TCTCCCTCATCGGCTCGGCGCGGGCTACCGTGTGCGCCAGCACCCCGACGCCAAGA	1227
1149	Db	TCTCCCTCATCGGCTCGGCGCGGGCTACCGTGTGCGCCAGCACCCCGACGCCAAGA	1208
1228	Qy	TACTTGGGCCATCGGCTAAGCATCGCTCCACCTCGCTAGCTGCGCTTCTTCTCCA	1287
1209	Db	TACTTGGGCCATCGGCTAAGCATCGCTCCACCTCGCTAGCTGCGCTTCTTCTCCA	1268
1288	Qy	TCGGCTTGGCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCG	1347
1269	Db	TCGGCTTGGCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCG	1328
1348	Qy	CGTGGGCTGCTGCTCGGGGTGCGCGCAACCGGTACAGGGGGTCACTCTCATGA	1407
1329	Db	CGTGGGCTGCTGCTCGGGGTGCGCGCAACCGGTACAGGGGGTCACTCTCATGA	1388
1408	Qy	CTTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGCAGCTTCTCTCTACTCCGGCA	1467
1389	Db	CTTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGCAGCTTCTCTCTACTCCGGCA	1448
1468	Qy	TCGCGCGCTGCGCTGGGTGTTCTTTCACCTTACTTCCGAGACCCGCGCGGACGC	1527
1449	Db	TCGCGCGCTGCGCTGGGTGTTCTTTCACCTTACTTCCGAGACCCGCGCGGACGC	1508
1528	Qy	TGAGAGAGATGACCAAGCTGTTGCGGACACGCGCCGCTCGGAATCAGACGACCCAG	1587
1509	Db	TGAGAGAGATGACCAAGCTGTTGCGGACACGCGCCGCTCGGAATCAGACGACCCAG	1568
1588	Qy	CCAAGGAGAGAGAGAGGTGGAAATGCGCGCCACTAACTGATCAAACTAACCCCAAAATC	1647
1569	Db	CCAAGGAGAGAGAGAGGTGGAAATGCGCGCCACTAACTGATCAAACTAACCCCAAAATC	1628
1648	Qy	ACCAATCTTAAGGGTTTCTTCCAAAACGTTGTGTACTGGCTAGCTAGCAAGTAGT	1707
1629	Db	ACCAATCTTAAGGGTTTCTTCCAAAACGTTGTGTACTGGCTAGCTAGCAAGTAGT	1688
1708	Qy	AGCAGCAACGTGGAGAGATTTCGTGATTCGCGCTTGTGTGAGAGCAGCGCGCGGACGA	1767

Dd	1689	AGCAGCAACGTGGGAAGAATTCGCTGTATCCGGCGTTCCTGGAGAGCGAAGCGGCCGCGACGA	1748
Qy	1768	CAAAGCTGAGCTCCAGCTCGAGACTTCCTTTAAAAATCATCTTCCTCAAGTAGCATGGATTTTATT	1827
Dd	1749	CAAAGCTGAGCTCCAGCTCGAGACTTCCTTTAAAAATCATCTTCCTCAAGTAGCATGGATTTTATT	1808
Qy	1828	TGCTCTTTGCTTTGCTCCGTAAAGTGTTACTATGCATGAAGTAATACCAGTAGTGTAGCAA	1887
Dd	1809	TGCTCTTTGCTTTGCTCCGTAAAGTGTTACTATGCATGAAGTAATACCAGTAGTGTAGCAA	1868
Qy	1888	GGCTGAGGTTGTGTGTAGCTACTAGAAAGTGTCACTCACCTTGTTCTTGTAAGAAAATGTTTT	1947
Dd	1869	GGCTGAGGTTGTGTGTAGCTACTAGAAAGTGTCACTCACCTTGTTCTTGTAAGAAAATGTTTT	1928
Qy	1948	AACGTGTAATTAAGCAGTATTTGTTGCAGTAATC	1980
Dd	1929	AACGTGTAATTAAGCAGTATTTGTTGCAGTAATC	1961

AK059423 1962 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:001-027-E03, full insert sequence.

ACCESSION AK059423
VERSION GI:32969441
KEYWORDS FLI_CDNA; oligo-capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
SpERMatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Brhartoideae; Oryzeae; Oryza.

REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team,
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yasaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Koijima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Ootomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Nariwakawa,R., Suyama,A., Mizuno,K., Yokonizo,S., Niihara,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata.K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki.Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12865764

REFERENCE 2 (bases 1 to 1962)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kawagawa,I., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Koijima,K., Koijima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara.C., Kurosaki.T., Kusumegi.T., Li.C., Lu.M.,
Masuda.H., Matsubara.K., Matsuyama.T., Miura.J., Miyazaki.A.,
Mizuno.K., Murakami.K., Murata.M., Nagata.T., Nakamura.M.,
Namiki.T., Nariakawa.R., Niikura.J., Nishi.K., Nomura.K.,
Numasaki.R., Ohneda.E., Ohno.M., Ohtsuka.K., Oka.M., Ooka.H.,
Osakazu.N., Ota.Y., Ootomo.Y., Ryu.R., Saitoh.H., Sakai.C., Sakai,K.,
Sakazune,N., Sano,H., Sasaki.D., Sato,K., Satoh,K., Shibata.K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano.S.,
Suqiyana.A., Suzuki.K., Suzuki.Y., Tagami.M., Tagami-Takeda.Y.,
Tagawa,A., Takemoto,A., Tanaka,A., Teraoka,H., Tokita,K.,

Tegawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
Tel: 81-29-838-7007, Fax: 81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL: <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Iehibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kuroaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, K., Sugano, S.,
Yoshiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ootato, N.,
Ota, Y., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

FEATURES
Location/Qualifiers
1..1962
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="001-027-E03"

ORIGIN

Query Match 93.2%; Score 1880.6; DB 8; Length 1962;
Best Local Similarity 97.6%; Pred. No. 1.3e-246;
Matches 1907; Conservative 1; Mismatches 45; Indels 0; Gaps 0;
28 GCCTTACACTCGACCGCCACTACTGTACACGCCCGCAGAGCGAGCTCTCTCTCTCTGCAC 87
DB 10 GCGAGCTCTCGACCGCCACTACTGTACACGCCCGCAGAGCGAGCTCTCTCTCTCTGCAC 69
QY 88 CACCGAGAGTGGCTTCGCGCGCGTGCAGGAGCGCGTGCAGGAGCGCGTGCAGGAGCGCG 147
DB 70 CACCGAGAGTGGCTTCGCGCGCGTGCAGGAGCGCGTGCAGGAGCGCGTGCAGGAGCGCG 129
QY 148 TCCGGTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 207
DB 130 TCCGGTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 189
QY 208 ATATCGGGGTGATGAGCGGGCGTGCCTGTACATCAAGAAGGACTTCAACATCAGTGACG 267
DB 190 ATATCGGGGTGATGAGCGGGCGTGCCTGTACATCAAGAAGGACTTCAACATCAGTGACG 249
QY 268 GGAAGGTGGAGTTCTCATGGGCACTACTGAACCTTACTTACTTACTTACTTACTTACTT 327
DB 250 GGAAGGTGGAGTTCTCATGGGCACTACTGAACCTTACTTACTTACTTACTTACTTACTT 309
QY 328 CGGGGCGGACGTGGAGTGGCGCGGGTACACCATCGTGTTCGCGCGCGTTCATAT 387

DB 310 CGGGGCGGACGTGCGAGCTGGATCGCGCGGGGGTACACCATCGTGTTCGCGCGCGTTCATAT 369
QY 388 TCTTCGCGGGGGGTTCCTCATGGGTTCGCGGTCAACTAGCCCATGTCTATGTTCGGCC 447
DB 370 TCTTCGCGGGGGGTTCCTCATGGGTTCGCGGTCAACTAGCCCATGTCTATGTTCGGCC 429
QY 448 GCTTCGTGGCGGATCGCGGTGGGTACGGCTCATGATCGCGGGGTACACCGCG 507
DB 430 GCTTCGTGGCGGATCGCGGTGGGTACGGCTCATGATCGCGGGGTACACCGCG 489
QY 508 AGGTGTGCGCGCGTTCGCGGTGGGTTCCTGACCTGCTCCCGGAGGTGTTCATCAACT 567
DB 490 AGGTGTGCGCGCGTTCGCGGTGGGTTCCTGACCTGCTCCCGGAGGTGTTCATCAACT 549
QY 568 TCGGATCTGCTCGGTACGCTCGAATATGCTTTCTCCGCTTCGCGGTGACCTTCG 627
DB 550 TCGGATCTGCTCGGTACGCTCTGAAATATGCTTTCTCCGCTTCGCGGTGAACTTCG 609
QY 628 GTTGGCGATCATGCTCGCATCGCGGGCGCTCGTGTCTGCTCGCTTCATGCTGC 687
DB 610 GTTGGCGATCATGCTCGCATCGCGGGCGCTCGTGTCTGCTCGCTTCATGCTGC 669
QY 688 TCGGATCTGCGAGTTCGCGGTGGGTTCATGAAGGAGCGCTTCGCGGACGCAAG 747
DB 670 TCGGATCTGCGAGTTCGCGGTGGGTTCATGAAGGAGCGCTTCGCGGACGCAAG 729
QY 748 TGGTGTGGAGAGACCTCCGACACGCGGGAGGAGCGCGGAGCGCTTCGCGGACATCA 807
DB 730 TGGTGTGGAGAGACCTCCGACACGCGGGAGGAGCGCGGAGCGCTTCGCGGACATCA 789
QY 808 AGCGCGCGCGCGCATCTCCGAGAGCTCGAGGAGCTCGAGGAGCGCTTCGCGGAGAG 867
DB 790 AGCGCGCGCGCGCATCTCCGAGAGCTCGAGGAGCTCGAGGAGCGCTTCGCGGAGAG 849
QY 868 GAGCGGAGAGAGAGCGGTTCGAGGAGCTCATCTGCTCCCGACCCCGCGCATGC 927
DB 850 GAGCGGAGAGAGAGCGGTTCGAGGAGCTCATCTGCTCCCGACCCCGCGCATGC 909
QY 928 GCGCATCTGCTGCTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 987
DB 910 GCGCATCTGCTGCTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 969
QY 988 CGGTGCTTCTACAGCGCTCTGCTGTTCAAGAGCGCGGATTAACGAGCGCAACACT 1047
DB 970 CGGTGCTTCTACAGCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1029
QY 1048 TCTTCGCGCACCACTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1107
DB 1030 TCTTCGCGCACCACTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1089
QY 1108 TCTTCATCGACCGCGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1167
DB 1090 TCTTCATCGACCGCGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1149
QY 1168 TCTTCATCGACCGCGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1227
DB 1150 TCTTCATCGACCGCGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1209
QY 1228 TACCTTCGCGCACCACTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1287
DB 1210 TACCTTCGCGCACCACTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1269
QY 1288 TCGGCTTCGCGCACCACTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1347
DB 1270 TCGGCTTCGCGCACCACTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1329
QY 1348 CGCTGGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1407
DB 1330 CGCTGGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1389
QY 1408 CTTCTGCTGCTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1467
DB 1390 CTTCTGCTGCTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 1449


```
/product="putative sorbitol transporter"
/protein_id="BAC83310.1"
/db_xref="GI:34393630"
/translation="MASAAPEAVAPKKGNVPFAFACILASMTSILLGYDGVMSG
ASLYIKDKVNIISDKVNIISLYLIGSFRAGRTSDWIGRRTYIVFAAIVFFAGA
FLMGAVNYAMFVAGTGVGYALMIAPVTAEVSPASARGFLTSFPEVFIQFI.
LLGVSNVAFSRLPLNIGIMLIGIAAPVLLALMLVGMPESPRLVMKGRILADAKV
VLEKTSITABEAERLADIKAAAGIPBELDGDVVTVPKRSGNEKRWKELILSPPA
MERILLSGIGHFFQASGIDSVLYSPRVFKSGAGITDDKHLGTTCAVGVTKLIFIL
VATFPLDRVRRPLLLSTSGMILSLGLGAGLTVQGHEDAKIPWAGIUSIATILAY
VAFPSIGLALTTWYSEIPIFLOVRALGCSLGAANRVTSGVISMTFLSLSKAITTGG
SFFLYSGIALTAWFFYTYLPTREGRTRLEEMSKLFGDTAAASEDEPAKEKKVEMAA
TN"
```

complement (7647..9324)

/gene="OJ1301.C12.4"

/complement(join(<7647..9067,9210..>9324))

/note="start and end point are not identified"

/complement(join(7647..9067,9210..9324))

/gene="OJ1301.C12.4"

/codon_start=1

/product="putative sorbitol transporter"

/protein_id="BAC83311.1"

/db_xref="GI:34393631"

```
/translation="MPSESTKLTAVDPKKNASYPFTCAILASMASIILGYDGVMS
GASLYIKDKITDQVEIILGILNIYSLVGSFAAGRTADWIGRRTYVFAAIVFFAR
ALLMGFGSDYATLMVGRFVAGVGVYAIMAPVTAISPASSRGFLTSFPEVSIINLG
ILGVSNVAFARPLSLGWRVLMGVCAAPVLLALMLVGMPESPRLVMKGRILADAK
AVLEKIDATPEEAERLADIKAAAGIPDDLDGDVVTVSKRSGNEKRWKELILSPPA
AMRIVLVAAGLHFFQASGVDVSVLYSPRVFKSGAGITDDKHLGTTCAVGVTKLIFIL
LVAFPLDRVRRPLLLSTSGMILSLGLGAGLTVQGHEDAKIPWAGIUSIATILAY
VAFPSIGLALTTWYSEIPIFLOVRALGCSLGAANRVTSGVISMTFLSLSKAITTGG
GGSFLLYAAISSLAWFFTRLPETRGTRLEEMSKLFGDTAAASEDEPAKEKKVEMAA
TSP"
```

11758..12731

/gene="OJ1301.C12.5"

/join(<11758..11894,11990..12071,12243..>12731)

/gene="OJ1301.C12.5"

/note="start and end point are not identified"

/join(11758..11894,11990..12071,12243..12731)

/gene="OJ1301.C12.5"

/note="predicted by FGENSEH etc."

/codon_start=1

/product="hypothetical protein"

/protein_id="BAC83312.1"

/db_xref="GI:34393632"

```
/translation="MAKRAETLGDHLRSLSSSSNOAPSWGAKRAGSALRDRSLPDS
EWDWSSIGNGPALVLSERVLDGVDVADLPELGRHVFVPTASGLVLLDRATFA
VRLNPLTRCLTDLRPATTVLDPTSLMISPVHLLHVLTVHGRGLDAGSATVYVEEN
ARKLAAPGDERWTVVEPDSPASVSSPAGRYFCATRAIRPVIVEVTRIDHRLPFPQ
LVVAKLAGVPSFPRNC"
```

/join(12995..17652,18451..19187)

/gene="OJ1301.C12.6"

12995..13731

/gene="OJ1301.C12.6"

/note="5' Terminal repeat"

/rpt_type=terminal

13837..17652

/gene="OJ1301.C12.6"

/note="probably inactive due to including stop codon(s) in

CDS

/pseudo

18451..19187

/gene="OJ1301.C12.6"

/note="3' Terminal repeat"

/rpt_type=terminal

complement(19137..20909)

/gene="OJ1301.C12.7"

complement(19137..20909)

/gene="OJ1301.C12.7"

/note="supported by full-length cDNA(s): AK110773"

19581..21159

/gene="OJ1301.C12.7"

19581..21159

/gene="OJ1301.C12.7"

19581..21159

/gene="OJ1301.C12.7"

19581..21159

/gene="OJ1301.C12.7"

19581..21159

```
Query Match 88.4%; Score 1783; DB 8; Length 25270;
Best Local Similarity 93.1%; Pred. No. 2.3e-233;
Matches 1916; Conservative 1; Mismatches 41; Indels 99; Gaps 1;
QY 28 GCTTACACTCGACCGCCTACTGTACACGGCCAGAGCGAGCCCTCTCTCTCTGCAC 87
DB 5129 GCGAGCTCTCGACCGCCACTACTGTACACGGCCAGAGCGAGCCCTCTCTCTCTGCAC 5188
QY 98 CACCGGAGATGCTTCCGCGCGCTCCGCGGCGCTCGCGCGAGCGCGTCCGCGGAGGCAACG 147
DB 5189 CACCGGAGATGCTTCCGCGCGCTCCGCGGCGCTCGCGCGAGCGCGTCCGCGGAGGCAACG 5248
QY 148 TCGGCTTCGCTTCGCGCGCTCCGCGGCGCTCCGCGGCGCTCCGCGGAGGCAACG 206
DB 5249 TCGGCTTCGCTTCGCGCGCTCCGCGGCGCTCCGCGGCGCTCCGCGGAGGCAACG 5308
QY 207 -----
DB 5309 GTAACATAACAAAGCGCATTCATCTCCGCGCGCGAGCAGCAGCTGCTGCTTCTCCCG 5368
QY 207 -----GATATCGGGGTGATGAGCGGGG 228
DB 5369 AGCATGGGCGCTCATCGCGATGTGCATGTGACATTCGAGATATCGGGGTGATGAGCGGGG 5428
QY 229 CGTCTGTACATCAAGAGGACTTCAACATCAGTGACGGGAGGTGAGAGTTCATCGG 288
DB 5429 CGTCTGTACATCAAGAGGACTTCAACATCAGTGACGGGAGGTGAGAGTTCATCGG 5488
QY 289 GCATCTGAACCTCTACTCGCTCATCGGCTCCTTCGCGCGCGGCGGCGGCGTCCGACTGGA 348
DB 5489 GCATCTGAACCTCTACTCGCTCATCGGCTCCTTCGCGCGCGGCGGCGGCGTCCGACTGGA 5548
QY 349 TCGCGCGGCGGTACACCATCGTTCGCGCGCGCTCATATTCTTCGCGGCGGCGGTCCTCA 408
DB 5549 TCGCGCGGCGGTACACCATCGTTCGCGCGCGCTCATATTCTTCGCGGCGGCGGTCCTCA 5608
QY 409 TCGGCTTCGCGCGTCAACTACGCGCATCTCATGTTTCGCGCGGCTTCGTCGCGCGGCGG 468
DB 5609 TCGGCTTCGCGCGTCAACTACGCGCATCTCATGTTTCGCGCGGCTTCGTCGCGCGGCGG 5668
QY 469 TCGGCTTCGCGCGTCAACTACGCGCATCTCATGTTTCGCGCGGCTTCGTCGCGCGGCGG 5728
DB 5669 TCGGCTTCGCGCGTCAACTACGCGCATCTCATGTTTCGCGCGGCTTCGTCGCGCGGCGG 5788
QY 529 GTGGCTTCCTGACGTCGTTCCGCGAGGTGTTTCATCACTTCGCGCATCTCTGTCGCGGTACG 588
DB 5729 GTGGCTTCCTGACGTCGTTCCGCGAGGTGTTTCATCACTTCGCGCATCTCTGTCGCGGTACG 5788
QY 589 TCTCGAATATGCTTTCTCCCGCTTCGCGCTGAACTCGGCTGCGGCAATCATGTCGCGCA 648
DB 5789 TCTCGAATATGCTTTCTCCCGCTTCGCGCTGAACTCGGCTGCGGCAATCATGTCGCGCA 5848
```


125 AACGTGAGGTTTCGCTTCGCTGCGCCATCTCGCTCCATGACCTCCATCCCTCCTCGGC 184
126 |||||
204 TACGATATCGGGGTGATGAGCGGGCGTTCGCTGTACATCAAGAAGAGATTCACATCAGT 263
205 |||||
185 TACGATATCGGGGTGATGAGCGGAGCGTTCGCTGTACATCAAGAAGAGATTCACATCAGT 244
245 |||||
264 GACGGGAAGGTGAGGTTCTCATGGGCATATCTGAACCTCTACTCGCTCATCGGCTCCTTC 323
265 |||||
245 GACACCCAGCTGGAGGTCCTCATGGGCATCTCAACGTGTACTCGCTCATTTGGCTCCTTC 304
324 CGCGCGGGGGAGAGCTCGGAGTGAATCGCGCGGGGTACACATCTGTGTTTCGCGCGGTC 383
325 |||||
305 CGCGCGGGGGAGAGCTCGGAGTGAATCGCGCGGGGTACACATCTGTGTTTCGCGCGGTC 364
364 |||||
384 ATATTCTTCGCGGGGGTTCCTCATGGGTTTCGCTCAACTAGCGCATGCTCATGTTTC 443
385 |||||
365 ATCTTCTTCGCGGGCGCTCATATGGGTTTCGCTCAACTAGCGCATGCTCATGTTTC 424
444 GGCGGCTTCGTCGCGGCATCGGCTGCGGTACGCTCATGATCGCGCGGGGTACAC 503
445 |||||
425 GGGCGCTTCGTCGCGGCATCGGCTGCGGTACGCTCATGATCGCGCGGGGTACAC 484
504 GCGAGGTGTCGCGCGGTTCGCGGTTCGCTCAACTAGCGCATGCTCATGTTTC 563
505 |||||
485 GCGAGGTGTCGCGCGGTTCGCGGTTCGCTCAACTAGCGCATGCTCATGTTTC 544
564 AACTTCGCGCATCTGCTCGGCTAGCTCTCGAATATGCTTTCTCCCGCTTCGCGCTGAAC 623
565 |||||
545 AACTTCGCGCATCTCTCGGATATGCTCCAACTTCGCGCTTCGCGCTTCGCTCCG 604
624 CTGGGTGGGCATCATGCTCGGATCGCGCGGGCGCTCGCTGTGCTCGGCTCATG 683
625 |||||
605 CTGGGTGGGCATCATGCTCGGATAGCGCGGGTCCGCTCGCTCGCTCGGCTCATG 664
684 GTGCTCGCGCATGCGCGGTTCGCTCATGAAGGAGCGCTTCGCGGAGCGC 743
685 |||||
724 GTGCTCGCGCATGCGCGGTTCGCTCATGAAGGAGCGCTTCGCGGAGCGC 724
744 AAGTGTGTGAGAGAGAGCTTCGACACGCGGAGAGCGCGGAGCGCTTCGCGGAGC 803
745 |||||
725 AAGTGTGTGAGAGAGAGCTTCGACACGCGGAGAGCGCGGAGCGCTTCGCGGAGC 784
804 ATCAAGCGCGCGCGCATCTCTGAGAGCTCGAGCGGAGCGCTTCGCGGAGCGC --- 860
785 ATTAAGACTGCGCGCGCATCTCTCTGCGGCTCGAGCGGAGCGCTTCGCGGAGCGC 844
861 AAGAGGAGCGGAGAGAGAGCGGCTTCGAGAGGAGCTCATCTGTCGCGGAGCGCGC 920
845 AACAAAGGAGAGAGAGAGAGAGCGGCTTCGAGAGGAGCGCTCATCTGTCACCGACCA 904
921 GCCATGCGGCGATCTCTGCTGCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGG 980
905 GCCATGCGGCGATCTCTGCTGCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGG 964
981 ATTCACCTTCGCTTCCTACAGCGCTTCGCTGCTTCAGAGCGCGCGGATTAAGAGAGC 1040
965 ATCGACGCGCTGCTGCTTCACAGCGCGCTAGTTCAGAGCGCGCGGATTAAGAGAGC 1024
1041 AACACCTTCCTTCGCGCATCTGCGCTTCGCTGCTACCAAGAGGCTTCATCTGTTG 1100
1025 AGCGCTTCGCGGCGACACCGTTCGCGGCTTCGCGGCGGCGGCGGCGGCGGCGGCGG 1084
1101 GCGACTTCCTTCATCGAGCGGCTGCGGCGGCGGCGGCGGCTGCTGCGGCGAGCGGCGG 1160
1085 GCGACTTCCTTCATCGAGCGGCTGCGGCGGCGGCGGCGGCTGCTGCGGCGAGCGGCGG 1144
1161 ATATCTCTCTCTCATCGGCTTCGCGCGGCGGCGGCTACCGCTCGCTCGGCGAGCGCGG 1220
1145 ATGCTGCTCTCTCTAGTGGGCTTCGCGAGCGGCGGCTACCGCTCATAGCGCGCGGCG 1204
1221 GCCAAGATACCTTCGCGGCGGCTGAGGATCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTTC 1280

1205 GAGAAATCACCTGGGCCATCGTCTGTGCATCTTTTGCATCGATCGCTACGTGGCTTC 1264
1281 TTCTCCATCGGCTTCGGCCCATCATCGTGGGTGTACAGCTCGAGATCTTCGCGCTCCAG 1340
1265 TTCTCCATCGGCTTCGGCCCATCATCGTGGGTGTACAGCTCGAGATCTTCGCGCTCCAG 1324
1341 GTGCGCGCTTCGGGTTCGCTCGCGGTTCGCGGCAACCGGCTCACAGCGGCTCATC 1400
1325 GTGCGCGCTTCGGGTTCGCTCGCGGTTCGCGGCAACCGGCTCACAGCGGCTCATC 1384
1401 TCATGACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1460
1385 TCATGACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1444
1461 TCGGGATCGCGGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1520
1445 GCGGATCGCTTCATTCGATGGGTGCTTCCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1504
1521 CGAGCTGAGGATGAGCAAGCTTCGCGGAGAGCGGCGGCTTCGGAATCAAG 1580
1505 CGAGCTGAGGATGAGCAAGCTTCGCGGAGAGCGGCGGCTTCGGAATCAAG 1564
1581 GAGCCAGCAAGGAG 1640
1565 GCGGAGCGGAG 1624
1641 CAAATACCAATCTTAAGGCTTTTCTGCAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
1625 CAAATGCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1684
1701 A-----AGTAGTAG 1745
1685 ACAAAGCGGATGAG 1744
1746 GGAGAGCGGAGCGGCGGCG 1762
1745 GGAGGCGAGCGCAAGC 1761

RESULT 8

AK119464

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone:001-133-F04, full

insert sequence.

ACCESSION

AK119464

VERSION

AK119464.1

GI:37989087

KEYWORDS

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

SOURCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1

AUTHORS

The Rice Full-length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-length cDNA Project Team;

Kikuchi, S., Satoh, K., Nagata, T., Kawagashita, N., Doi, K.,

Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,

Kojima, K., Naniki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,

Ohtsuka, K., Shisshiki, T., Foundation of Advancement of International

Science Genome Sequencing & Analysis Group, Otsu, Y., Murakami, K.,

Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

Kurotsaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,

Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,

Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,

Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Riken,

Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,

Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,

Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

Science 301 (5631), 376-379 (2003)

JOURNAL

QY 1599 AAGAAGTGAATGCCGCCCTAACTGATCAA 1632
 Db 1628 AAGCGAAAGTAGGGAGATGAACAGTGTGAGCTA 1661

RESULT 11
 AP004185/c 25270 bp DNA linear PLN 02-SEP-2003
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
 BAC clone: OJ1301_C12;
 ACCESSION AP004185
 VERSION AP004185.3 GI:32261236
 KEYWORDS
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 clone: OJ1301_C12
 JOURNAL Published Only in Database (2001)
 REFERENCE 2 (bases 1 to 25270)
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Direct Submission
 JOURNAL Submitted (20-SEP-2001) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program; Kamonndai
 1-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 COMMENT On Jun 25, 2003 this sequence version replaced GI:21280394.
 Genes were predicted from the integrated results of the following:
 GENSCAN (http://ccr-081.mit.edu/GENSCAN.html), FGENESH
 (http://www.softberry.com/), GeneMark.hmm
 (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
 (http://opal.biology.gatech.edu/glimmer/glmr form.html), RiceHMM
 (http://rgp.dna.affrc.go.jp/RiceHMM/), SplinePredictor
 (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI Nonredundant Protein
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
 sequence database at RGP or DDBJ. Protein homologies of the coding
 regions were searched against NCBI Nonredundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative', and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as
 a probable 'hypothetical' protein and is included as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from -21M13 to M13rev of the BAC
 clone. This is a partial sequence of OJ1301_C12 clone. This sequence
 of OJ1301_C12 clone has an overlap with P0453603 clone (DDBJ:
 AP004276) at 5' end and an overlap with OJ1127_E01 (DDBJ: AP003747)
 at 3' end. The sequence was generated by combining Monsanto and
 RGP-Japan sequencing data. Detailed information on overlap and
 assembly quality together with annotation of this entry is
 available at
 http://rgp.dna.affrc.go.jp/GenomeSeq.html.
 Location/Qualifiers
 1. .25270
 /organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="7"
 /clone="OJ1301_C12"
 779. .1103
 /gene="OJ1301_C12.1"
 join(<779. .940,1011. .>1103)
 /gene="OJ1301_C12.1"
 /note="start and end point are not identified"
 join(779. .940,1011. .1103)
 /gene="OJ1301_C12.1"
 /note="predicted by GENSCAN etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAC83309.1"
 /db_xref="GI:34393629"
 /translation="MYRERAQRGRSSATVFLASSRQLGKGGKVFARDGNPAGCGPY
 CRPASRATLRLSADMTTAACDGSSAAVDHSGFRFS"
 complement(join(1344. .1480,2498. .2686,3144. .3198))
 /gene="OJ1301_C12.2"
 complement(join(1344. .1480,2498. .2686,3144. .3198))
 /gene="OJ1301_C12.2"
 /note="hypothetical ORF
 predicted by GENSCAN
 this category is not included in IRGSP standard"
 5120. .7180
 /gene="OJ1301_C12.3"
 join(5120. .5308,5408. .7180)
 /gene="OJ1301_C12.3"
 /note="supported by full-length cDNA (s): AK059423"
 join(5121. .5308,5408. .7180)
 /gene="OJ1301_C12.3"
 /note="supported by full-length cDNA (s): AK102676"
 join(5197. .5308,5408. .6828)
 /gene="OJ1301_C12.3"
 /note="contains EST (s):
 C20132(EI1563), AU163952(EI1563), AU037469(S5067),
 AU033198(S5067)
 contains full-length cDNA (s): AK059423, AK102676"
 /codon_start=1
 /product="putative sorbitol transporter"
 /protein_id="BAC83310.1"
 /db_xref="GI:34393630"
 /translation="MASAALPEAVAPKKNVRFAPACAILASMTSILLYDIGVMSSG
 ASLYIKKDFNIGDKVEVLGMILNLSLIGSFAAGRTSDWIGRRYTVFAAVIPFAGA
 FLMGFAYNAMLGFGVAGVYALMTAPVTVAEVSPASARGFTLTSPTVFINDGI
 LLGVSNYASRLPLNLGIMLIGIAPSVLLALMLVGMPSRPLVWKGRLADAKY
 VLEKTSABEAERLADINKAAGIEEELDGVTVTVPKGSGNEKRWKLELISPTPA
 MRLLESGIHFFQASGIDSVLYSPRVFSAGITDDKHLGTTCAVGTUFIIL
 VATPFLDRGRPLLSSTGCMILSLIGLGLTVVQHPDKIMPAIGLSASTLAY
 VAFPSILGIGITWYSEIPEPQVRALGSLGVANRVTSVIMSTFLSKAITGG
 SFLYSIGTIALAWVFYVYLPETRGRTLEEMSKLPGDTAAASESDPAKEKKKVEVAA
 TN"
 complement(7647. .9324)
 /gene="OJ1301_C12.4"
 complement(join(<7647. .9067,9210. .>9324))
 /gene="OJ1301_C12.4"
 /note="start and end point are not identified"
 complement(join(7647. .9067,9210. .9324))
 /gene="OJ1301_C12.4"
 /codon_start=1
 /product="putative sorbitol transporter"
 /protein_id="BAC83311.1"
 /db_xref="GI:34393631"
 /translation="MPESYKLTAMAVDPKKKNASVAFACAILASMASIILGIDIGVMS
 GASLYIKKDLKITDVOVEILGNIYLSVGSAGRTADWIGRRYTVFAAVIPFAGA
 ALLMGFSDYATLMVGRVAGVGVAIMIAPVYTAETI SPASSRGFTLTSPTVFINDGI
 ILIGVSNYAFARLPLSLGHRNMLVGCAAPSVLLALMLGMPESRPLVWKGRLADAK
 AVLEKTIADTPEASERLADINKAAGIPDDLDGVTVTVSKRGGRGCQVWRELVSPTF
 AMRIRVLAAYHFFQASGIDSVLYSPRVFSAGITDDKHLGTTCAVGTUFIIL
 LVAFLLDRGRPLLSSTGCMVFLVGLATGLTVYGGSPDAQVPSVAVGLCVASIL
 AYVAFFSVGLGPMGVSIVTSIFPLRARALGFAVAVACNRTSVGIVSMSTFLSKAITT
 TN"

FEATURES
 source
 1. .25270
 /organism="Oryza sativa (japonica cultivar-group)"


```

RESULT 13
AP004365/c
LOCUS
DEFINITION
  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
  PAC clone:P0458E05.
ACCESSION
  AP004365 BAC00010
VERSION
  AP004365.3 GI:21902053
KEYWORDS
  Oryza sativa (japonica cultivar-group)
SOURCE
  Oryza sativa (japonica cultivar-group)
  Oryza sativa (japonica cultivar-group)
  Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
  1
  Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T.,
  Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y.,
  Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K.,
  Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K.,
  Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y.,
  Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Itoh, S., Itoh, T.,
  Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S.,
  Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T.,
  Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M.,
  Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M.,
  Ohka, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T.,
  Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Teuji, K., Waki, K.,
  Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K.,
  Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.I., Eun, M.Y.,
  Yano, M., Jiang, J. and Gojobori, T.
  The genome sequence and structure of rice chromosome 1
  Nature 420 (6913), 312-316 (2002)
  22337376
  12447438
  2 (bases 1 to 193577)
  Sasaki, T., Matsumoto, T. and Yamamoto, K.
  Direct Submission
  Submitted (19-NOV-2001) Takui Sasaki, National Institute of
  Agricultural Sciences, Rice Genome Research Program, Kamondai
  2-1-2, Teukuba, Ibaraki 305-8602, Japan
  (E-mail: tsasakionias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
  Tel: 81-298-38-7441, Fax: 81-298-38-7468)
  On Jul 17, 2002 this sequence version replaced gi:19773518.
  Genes were predicted from the integrated results of the following:
  GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
  (October 1998 version). The genomic sequence was searched against
  NCBI Nonredundant protein database, nr
  (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
  RGP. Protein homologies of the coding regions were searched against
  NCBI Nonredundant protein database with BLASTP2.0. ESTs represent
  the identified cDNA sequences using BLASTN 2.0 with the
  corresponding DBJ accession no. and RGP clone ID.
  A gene with identity or significant homology to a protein is
  classified based on the protein name to indicate the homology level
  such as same name, 'putative,' and '-like protein'. A gene without
  significant homology to any protein but with EST homology (covering
  almost the entire length of partial sequence) is classified as an
  'unknown' protein. A gene predicted with a gene prediction program
  is classified as a 'hypothetical' protein.
  The orientation of the sequence is from T7 to SP6 of the PAC clone.
  This sequence of P0458E05 clone has an overlap with P0483G10 (DBJ:
  AP003263) at the position 1 to 50,738 of 5' end and an overlap with
  OJ1656 All (DBJ: AP003448) at the position 188,119 to 193,577 of
  3' end. Detailed information on overlap and assembly quality
  together with annotation of this entry is available at
  http://rgp.dna.affrc.go.jp/GenomeSeq.html.
  Location/Qualifiers
    1..193577
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="genomic DNA"
    /cultivar="Nipponbare"
    /db_xref="taxon:39947"
    /chromosome="1"

/clone="P0458E05"
join(1748..1991,2570..2750,2997..3243,5685..5813,
5922..6140,6199..6213)
/gene="P0458E05.1"
join(1748..1991,2570..2750,2997..3243,5685..5813,
5922..6140,6199..6213)
/gene="P0458E05.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05602.1"
/db_xref="GI:21902054"
/translation="MRRLFRPLRPMGAVARRSRWRVPAVTAVYGHARRAECVRCG
AKMGCKKWEKGGGRARGFGFGEVSSVWVAVTAHVSYPVRAADGACBPPL
TGAHPSPFRRRRRPARARRRRRIIRIPPPRAIKELAPLSSLSVADLRLP
ALSARDLPPRLRLRLRIAADFVHLAELARSAIVPVPFSPRAAFRSGRLLC
RRRPGRCFIBFGEIRRESCGRREFITKYSPIEITPIYSLDLDHMKDRKVEEE
BAEMIOQMBLGLGESITDEYLACLNLPHPGPPRPUTKTKLQGEQLHDIYIRHTV
RLKYVKSIS"
7085..10860
/note="5' LTR"
12336..117054
/gene="P0458E05.2"
12336..117054
/gene="P0458E05.2"
/codon_start=1
/product="putative polyprotein"
/protein_id="BAC05603.1"
/db_xref="GI:21902055"
/translation="MASRPREAARAPARFGFEDGRVEEENHESANEQSHSNHSRHT
VSRNAEVPRRSRSHRTSVRDAEMGQSHRSHTASNGSEERLPFPFHVETNMQHL
NAVOTIQLGLTTAIASPOHNAHGHPHMGNNRSKLTDFLRSPFPFSQVPEVAD
DWLKDVKLVLCVCTPVEKMLYASHQLRGAADWENYCNAPPEPTNIAWDEPATAF
PAHVPESTIDMKKEEENELKQGNSSINEVLSQNKLAARYAPEVDTDKIRKFLKG
IAGWRLQLAHDFTFCHMINKALLLEDARKEATEEYKKSXHQGNSFGASRPYR
GPMQYHOSVTOANRPGYAPRPQINRPAQPOQRASGTAPNSVTSFKSPQSPASV
QCFRQNHGYARCPQPTNTNSHANGSTARTPTAAQSRPSSQASGGSASRAN
FGRGRVNIQAEATQADAPVVMFMSVNPVAILFDGSHSFISQAFVRNGMKTQ
NLRVPMIVHSPGRNIRATQICEVNLRIEEDVFLAKPILVDSLDILGMDWLAKH
QIDCAKXSITLQGGKQVFTPTNPTASSILTCLQVTSLESPVCEYVDFVPEP
LKWPPDPEIFBFIELAPKAPRYMAANELAEVKQIELESKGVYVRSSP
CAPVLVKKQSGRWIDYALNEVINKYPLPRIDDLFDQJGKARVSKILRSG
YHQKIRSEDPKTAFTRYGLYFTVNSGPIDNAPAFMMNMKIFNEHLDQVYVF
IDNLIYSKEEBAEHLRMEKLDHQLFAKSKCFEFLDRPMDRPAEVLHVISNGVEF
DPSKEAVLAWNPKNVSEIRSFGLAGYRFTIEGSKLAPMTLEKKEKQWSA
ACEFSOBLKRLTAPVLTPDKOPEIFCDASRQGLCMQEQKVVAYARQLR
PHENYPTHLEAAVHVALKWRHLYIGNRECVTQHSKLYFTOTELNMRQKWL
ELIYDYLGIHYHPCKANVADALSHKAYCNVAOIRPDQHLCRELEKRLTVQSGV
PASLIVQTLSEQREACKDEGKELIKRIQEKDNRSIDDKGTWCGPRICVPKAE
KELRLDILKEAHSAISHPGSTKMYQIKAFYFWAGMKRDUVAEYVALCDICQVKAE
HORFAGLLQPLPIPEWKWEEIGMDFITGLPTPSGYDSIWVLDRLTKSAHFVPKTT
YDQKLAELMYTHVYVRCFCPKIVSDRGTOFTSRFWKQLHEALGTDLNFTAYHPQT
DQOERVNQIILEDMLRACALDFEGTWDRCLYAEFSYNNYSYQASTQNSPNAEGRKC
RTPLCNWVGEALVFGPDLKSAEEOVKLIERLKTACNQKYNADNRRLDLEPEKGD
HYLVESPLRGMRBEGMSGLAPRYIGYLLITARHGEVAYOLEPEGLADVHNVFHVS
QLKCLRVPQEQVPLGNIENLKNITKRPKIVLEVERQTERTKITKEYKQVQNSHSE
DEATWERBELLRAEFPDLLP"
17098..20885
/note="3' LTR"
20886..24214
/note="5' LTR"
25856..30421
/gene="P0458E05.3"
25856..30421
/gene="P0458E05.3"
/codon_start=1
/product="putative polyprotein"
/protein_id="BAC05604.1"
/db_xref="GI:21902056"
/translation="MVVTRTSRATGCSNGSEERADGVHPNGDSNGPPLPENPLTA
QVMAHQTPMAMQMQQQQHQMQRMQHAEQHQFGFPFQSKLPELVRPPT
FSSTNTMEANDWLHAEIKNLLOCDQEKVAFATHQLOQGFASAMNDNWTAPPGT
EVTWAEFCRSPKQVPGVVAQKREFRALHQGNRTVTEYLHFNRLIARVAPDPRQ
DAEKQEKTFMAGLDDELTLNQLISGDYADFRLVDAIKAIQEDQQRNDRKRAAQFRSNG

```

GSQHRPRFTPGCGGPTTMTVROHRRPFPNPNPHQGTSGSQNHGGGPNRGAAPRPPMA
PAGSGPAQAKKFTGAKPGSCFNGELGHPADCKPKPRRAGPREFIGARVNHASABEAQ
AAEVLVLGPUNNISTPATVLPDSGATHSEISKFFVGMHGLISEBELSPMRVTPGNS
TSVQFSPSINIEQRSPFLANLILLESKOLDVLNGMDWLTKFKGVUDCARVTLTNE
KGTVEVLVSPPKGAUNQIETEPVDTKEKNRKLHEDIPICBYPFVFFEDLITM
PPREBEFIDFLAPGAPYKPYRMAANELAEVKQDEQLOKGYRSESTPFWGAPV
IFVEKDKTRMCDVYALNEVITIKKYLPLRIDDLDQKGAKFVKIDLRSGYHQL
RIEEDIDPKTAFTRVGLYEYVMSFGLTNAPAFMNLKNNKFMFELDFVVFVDDI
LIYSKBEHEHRLVLEKLKHEQLYAKFSKDFWLTEVFLGHVITAGQVAVDPFN
VESVTKPPTKQJRSFGLAGYVRRFENFSRIARDWTOLKKDEKFKWTABCDK
SPFELKRLVSAVILLPDQTKDFQNYCDASRHGLGCVLMQEGRVVAYASRQLRPHG
NYTHDELAALVHAUKIMRHYLIGRCEVITDTHKSLKVIPTQDNLNRQWLELILK
DYDMSHYHFGKANVADLSRSYCTALCIEGMCEBELRQBEHLNIGVIEHFVAAAL
EAREPTLVQVAAQVNDPEIAELKQMRVKARDFLEDEHGTIWMGERLCPDDKELK
DLILTEAHQPTPIHPSGTKNVQDLXKFKWVMSRREIAEFVALCNVQKVAEHORP
AGLLOPLQIPEWKWEIGMDFITGLPRTSGHDSIWWVDRLTKVAHFIPVHTVTKG
RLAELVLARLCHGVKPIVSDRSQFTSKFWKLOBELGTFELNESTAVHPTQDGT
ERNQILEDMRACALDFGAWDKSLPYAFBSYNNYSQASLQMAPEALYGRKRTPL
FWDQGERQFQTEVLAAEKEKRIIRERLRIASROKSYADNRRLBELTFEAGDYVYL
RVTPLRGVHRRFQTKLAPRFVGPYKILERRGEVAYOLELPSNMIGIHDFVHSQJLK
CLRPVBEQADSEHIDIQEDLTYVEKPVRIIDTSETRNRKNKVTFRVQVSHSEBEAT
WEREDELKAAHPHLETSSESRRGDSV"
30434..33762
/note="3' LTR"
complement(join(36988..37110,40188..40622))
/gene="P0458505.4"
complement(join(36988..37110,40188..40622))
/gene="P0458505.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05605.1"
/db_xref="GI:21902057"
translation="MDAAGTTKISDFGLAKLQPDQTRFTTSVGRTRYGLASEWVRGA
GPVMKADVSYGVVILETTVACRRSMEMEAEGERTELAEWAYEWLLVKSEAKSAMS
SDTEVLAAREVVRVVMQVQAEPOSRRSPSMGVILMLOGRLERQHHEBEAFAFGRL
CGBHLVHREDVGEFSEWQSHLVLI"
complement(40716..41015)
/note="3' LTR"
complement(41118..45086)
/gene="P0458505.5"
complement(41118..45086)
/gene="P0458505.5"
/note="probably inactive due to stop codon(s) in CDS
pseudogene, putative retroelement pol polyprotein"
/pseudo
complement(45135..45440)
/note="5' LTR"
join(56253..58166,58250..58411,58862..59407,59787..59867,
60240..60320)
/gene="P0458505.6"
join(56253..58166,58250..58411,58862..59407,59787..59867,
60240..60320)
/gene="P0458505.6"
/product="putative chromosome condensation factor"
/codon_start=1
/protein_id="BAC05606.1"
/db_xref="GI:21902058"

Query Match 45.1%; Score 909.2; DB 8; Length 193577;
Best Local Similarity 79.9%; Pred. No. 1.9e-114;
Matches 1096; Conservative 1; Mismatches 269; Indels 6; Gaps 2;
QY 219 ATGAGCGGGCGTGTATCATCAAGAAGACTTCAACATCAGTGACGGGAAGGTGGAG 278
Db 151179 ATATCGGGGGCGTGCCTGTTTCATCAAGAGAGACATGAAGATCAGCGACGTGAGGTGGAG 151120
QY GTTCTCATGGGCATACCTGAACCTCTACTCGGTCTCATCGGCTCTTCGCGGCGGGCGGACG 338
Db 151119 ATCTGTCTGGCATCTCAACCTCTACTCCCTCGTGGCTCTTCGCGCGCGCGGACG 151060
QY TCGAGCTGGATCGCGCGGCGGTACACCATCGTGTTCGCGCGCGGTCTATCTTCGCGGGG 398
Db 151059 TCCAGCTGGATCGCGCGCGCGCTCACCATCATCTCCCGCGGGTCATCTTCTTCGTCGGC 151000

399 GSGTTCTCATGGGTTTCGGCGTCAACTAGCCATGCTCATGTTTCGGCGGCTTCGTGGCC 458
Db 150999 GCATCATGATGGGCTCTCCGTCAACTACCCATGCTATGGCGCGCGCTTCGTGGCC 150940
QY 459 GGCATCGCGGTGGGTACCGCGTCAATGATCGCGCGCGGTGTACACCGCGAGGTTCGGCG 518
Db 150939 GGCATCGCGGTTCGGGTACCGCTTCAATGATCGCGCGCGGTCTACACCGCGAGGTTCGGCG 150880
QY 519 GCGTCGCGCGGTGGCTTCTTGACGTCGTTCGCGGAGGTGTTTCACTCAACTTCGGCATCTCG 578
Db 150879 GCGTCGTTCGGCGGCTTCTTCACTTCTTCGCGAGGTGTTTCACTCAACTTCGGCATCTCG 150820
QY 579 CTCGGGTACGTTCGAATATGCTTTTTCGCGTTCGCGCTGAACTCGGTTGGCGCATC 638
Db 150819 CTCGGGTACGTTCGAATATGCTTTTTCGCGTTCGCGCTGAACTCGGTTGGCGCATC 150760
QY 639 ATGCTTCGGCATCGCGCGCGCGCGCTTCGCTGCTCGCGCTCATGTTGCTCGGATCGG 698
Db 150759 ATGCTTCGGCGTTCGGCGCGCGCGCGCTTCGCTGCTCGCGCTCATGTTGCTCGGATCGG 150700
QY 699 GAGTCGCGCGGTGGCTGCTCATGAAGGAGCGCTTCGCGGACCGCAAGGTGCTGTTGAG 758
Db 150699 GAGTCGCGCGGTGGCTGCTCATGAAGGAGCGCTTCGCGGACCGCAAGGTGCTGTTGAG 150640
QY 759 AAGACTCTCGACACCGCGGAGGAGCGCGCGGAGCGCTTCGCGGACCATCAAGCGCCCGCC 818
Db 150639 GAGACTCTCGACACCGCGGAGGAGCGCGCGCGCTTCGCGGAAATCAAGGAGCGGTG 150580
QY 819 GGCATCCCTGAGGAGCTCGACGCGGACGCTGGTGACCGTCCCCAAGAGAGGAGCGGAAAC 878
Db 150579 GGCATCCCGCGCGACCTCGACGCGTGCCTGCGCGCTTCGCGAAGAGGCGCGCGG--C 150523
QY 879 GAGAGCGGGTGTGAAGAGAGCTCATCTGTTCGCGACCGCGCGCATCGCGGCGATCTG 938
Db 150522 GAGAGCGCGTGTGAAGAGAGCTCATCTGTTCGCGACCGCGCGCTTCGCGGCGGTGTC 150463
QY 939 CTGTCGGGATCGGATCCACTTCTTCAGCATGTTGGGATTCCTCCGTCGCTTC 998
Db 150462 CTGTCGGGATCGGATCCACTTCTTCAGCATGTTGGGATTCCTCCGTCGCTTC 150403
QY 999 TACAGCCCTCTCGTGTTCAGAGCGCGGATTAAGAAACGACAAACACTTCTTGGGACAC 1058
Db 150402 TACAGCGCGCGTGTTCAGAGCGCGGATTAAGAAACGACAAACGACGTCGTCGACG 150343
QY 1059 ACTTGCGCGTTCGCTGTACCAAGAGGCTTTTCATCTTGTGGCGACTTCTTCATCGAC 1118
Db 150342 ACTTGCGCGTTCGCGGTCAACCAAGACGCTTTCATCTTGTGGCGACTTCTTCATCGAC 150283
QY 1119 GCGTCGCGCGCGCGCGCTGTTGCTGGGACGCGCGGATTAATCTCTCTCTCATC 1178
Db 150282 GCGTCGCGCGCGCGCGCTGTTGCTGGGACGCGCGGATTAATCTCTCTCTCATC 150223
QY 1179 GCGTCGCGCGCGCGCTCAACGCTCGTCGCGACGCGCGGATTAATCTCTCTCTCATC 1238
Db 150222 ACCCTCGGCTTAGGCGCTCAACGCTCATCGGTGAGGACGCGCGG--GGCGGCTGGGCG 150166
QY 1239 ATGCGGCTAAGCATCGCTTCACCGCTTCGCTGCTGGGACGCGCGGATTAATCTCTCTCTCATC 1298
Db 150165 ATGCGGCTAAGCATCGCTTCACCGCTTCGCTGCTGGGACGCGCGGATTAATCTCTCTCTCATC 150106
QY 1299 CCATCAAGCGGCTTACAGCTCGGAGATTTTCGCGCTCCAGGTCGCGGCGGTGGGCTGC 1358
Db 150105 CCATCAAGCGGCTTACAGCTCGGAGATTTTCGCGCTCCAGGTCGCGGCGGTGGGCTGC 150046
QY 1359 TCGCTCGGCTCGCGCGCAACCGCGCTTCACGAGCGGCTCATCTCCATGACCTTCGCTGC 1418
Db 150045 GCTCTCGGCTCGGCTCAACCGCGCTTCACGAGCGGCTCATCTCCATGACCTTCGCTGC 149866
QY 1419 CTGTCAAGCGGCTTACAGCTCGGAGATTTTCGCGCTCCAGGTCGCGGCGGTGGGCTGC 1478
Db 149985 CTGTCAAGCGGCTTACAGCTCGGAGATTTTCGCGCTCCAGGTCGCGGCGGTGGGCTGC 149926

QY 1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCGCGCGGAGCGCTGGAGAGATG 1538
 Db 149925 GCTTGGCTCTTCTTCTTCACTACTCTCCCGAGACTGTTGGACGACGCTCGAGCAGATG 149866

QY 1539 ACCAGCTGTTCCGGGACACGGCCCGCTCGGAATCAGACGACGCA 1590
 Db 149865 GCGGAGCTCTTCCGCATCCACAATGCGCGGCGAGCAGCACTGTCGGCCA 149814

RESULT 14
 AC116426 155037 bp DNA linear PLN 28-MAR-2002
 Locus Genomic sequence for Oryza sativa, Nipponbare strain, clone
 DEFINITION OSUNB0058P18, from chromosome 3, complete sequence.
 AC116426
 ACCESSION AC116426.1 GI:19774446
 VERSION HTG.
 KEYWORDS
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 155037)
 AUTHORS McCombie, W.R., Spiegel, L., de la Bastide, M., Preston, R., Kuit, K.,
 Ferraro, K., Nascimento, L., Zutavern, T., Ballia, V., Bell, M.,
 Baker, J., Santos, L., Miller, B., Katzenberger, F., Muller, S.,
 King, L., Yang, C., Dike, S., O'Shaughnessy, A., Palmer, L. and
 Dedhia, N.
 TITLE Genomic sequence for Oryza sativa, Nipponbare strain, clone
 OSUNB0058P18, from chromosome 3, complete sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 155037)
 AUTHORS McCombie, W.R.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAR-2002) Lita Annenberg Hazen Genome Center, Cold
 Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
 NY 11724, USA
 COMMENT This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest.
 Clone OSUNB0058P18 overlaps finished clone OSJNBa0091J11
 (AC115687) from base 1 to base 94350. The overlap is from base
 86812 to base 18163 on OSJNBa0091J11.
 FEATURES
 source
 1..155037
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="3"
 /clone="OSUNB0058P18"
 /clone_lib="EcoRI"
 3770..3840
 misc_feature
 /note="We believe the assembly to be correct. The
 sequence is covered by a PCR product which was amplified
 with a high fidelity polymerase. There is additional
 coverage by a single subclone with quality below phred30."
 16142..16264
 misc_feature
 /note="We believe the assembly to be correct. The
 sequence is a polynucleotide (ATNACATn) repeat in which
 the exact number of copies is unknown. Two subclones span
 the repeat into unique sequence on both sides."
 65646..65667
 misc_feature
 /note="We believe the assembly to be correct. The
 sequence is a mononucleotide (C) repeat in which the exact
 number of Cs is unknown. One subclone in the region spans
 into unique sequence on both sides."
 136562..136740
 misc_feature
 /note="We believe the assembly to be correct. The

sequence is covered solely by direct dye terminator
 sequencing reactions done with BAC template DNA. The
 sequence is high quality."

ORIGIN

Query Match 37.6%; Score 759.4; DB 8; Length 155037;
 Best Local Similarity 73.0%; Pred. No. 4,8e-94;
 Matches 992; Conservative 1; Mismatches 357; Indels 9; Gaps 1;

QY 207 GATATCGGGGTGATAGCGGGGCGTCTGTACATCAAGAAGACTTCAACATCAAGTAC 266
 Db 85013 GACATCTCGGTGATGAGCGGCGCAGAGTTTCATGAAGAAGACCTGAACATCTCGAC 88954

QY 267 GGAAGGTGAGGTTCTCATGGGCACTGAACCTCTACTGCTCATGGCTCTCTTCGG 326
 Db 88953 GCGAAGGTGGAGTCTCTCGCGGCATCATCAACATCTACTGCTCGGTCTGCTCGCG 88894

QY 327 GCGGGCGGACGCTCGGACTTGGATCGCGCGGCTACACATCGTTCGCGCGCGTCA 386
 Db 88893 GCGGGCGGACGCTCGGACTTGGATCGCGCGGCTACACATCGTTCGCGCGTCA 88834

QY 387 TTCTTCGCGGGGSGTTCTCATGGGTTGCGCGTCAACTAGCCATCTCATGTTTCGGC 446
 Db 88833 TTCTTCGCGGGGCGTCTCATATGGGCTCTGCGCGAGTACCGATCTCATGTTTCGGC 88774

QY 447 CGCTTCGTTGGCGGCGATCGGCGTGGGCTAGCGGCTCATGTCGCGCGGCTACACGCC 506
 Db 88773 CGCTTCGTTGGCGGCGGTTGGGCTAGCGGCTCATGTCGCGCGGCTACACGCC 88714

QY 507 GAGGTGTTCGCGCGGCTCGCGGCTGCTTCCTGAGCTGTTCCCGGAGGTTTCATCAAC 566
 Db 88713 GAGGTGGCGCGACGCTCGCGCGGCTGCTTCCTGAGGTTTCATCAAC 88654

QY 567 TTCGGATCTCTCTCGGTAGCTCTCGAATATGCTTCTCCGCTTCGCGCTGACCTC 626
 Db 88653 GCGCGGCTCTCTCGGTAGCTCTCGAATATGCTTCTCCGCTTCGCGCTGACCTC 88594

QY 627 GGGTGGCGCATCATGCTCGGCATCGCGCGGCGGCTCGGTCTGCTCGCGCTCATGTG 686
 Db 88593 GACTGGCGGTATGTTCTTAGTCTGGCGCGGCTCGCGCTTCTCTCGCGTGGCGTC 88534

QY 687 CTCGCGATCGCGGATGCGCGGCTGCTGTCATGAAGGAGCGCTCGCGGAGCGCAAG 746
 Db 88533 CTCGCGATCGCGGATGCGCGGCTGCTGTCATGAAGGAGCGCTCGCGGAGCGCGCA 88474

QY 747 GTGGTGTGGAGAAGACCTCGGACACGCGGAGGAGCGCGGAGCGCTCGCGACATC 806
 Db 88473 CGTGTCTCGAGAAGACCTCGGACATCGCGCGGAGGAGCGGCTCGCGACATC 88414

QY 807 AAGGCGCGCGCGCATCCCTGAGGAGCTC-----GACGGCGAGCTGGTGAACCTC 857
 Db 88413 AAGAAGCGGTTCGGCATCCCCGAGAGGATCTCCGACGAGGACGAAGTGTAGCGCTC 88354

QY 858 CCAAGAGAGGAGCGGAAGAGCGGCTGAGAGGAGCTCATCTGTCCCGGAC 917
 Db 88353 CACAAGAGCGAGGGTCTCCACGCGAGGCTGCTGAGGAGGCTCTGCTCTCGCGCGGAC 88294

QY 918 CCGGCATTCGCGCGCATCTGCTGTCGCGGATCGGCATCCACTTCTTCCAGCATCGCTG 977
 Db 88293 CCGGCGGTGCGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 88234

QY 978 GGCATTACTCGTCTCTTCTACAGCCCTCTGCTGTTCAAGAGCCCGGATTAAAGNAC 1037
 Db 88233 GGCATTACTCGTCTCTTCTACAGCCCTCTGCTGTTCAAGAGCCCGGATTAAAGNAC 88174

QY 1038 GACAAACACTTCTTCGGGACCACTTGGCGGCTTGGCTGTGTCACCAAGAGGCTTTTTCATCTG 1097
 Db 88173 GACTCGGACTTCATCGGAGCTTCCGTGGGCTGGGCGCGCAGCAGAGCTCTTTCATCTC 88114

QY 1098 TTGGCGACTTCTTTCATCGAGCGCTCGGGCGGCGCGCTGCTGCTGGGAGCAGCGGC 1157
 Db 88113 GTGGCGACTTCTTCTTCGAGCGCTCGGGCGGAGCGGCTGCTCTCTCAGCGCGCGC 88054


```
QY 1158 GGGATAATCTCTCCCTCATCGGCTCGGCGCGGGCTACCGCTCGTGGCCAGCACCC 1217
Db 88053 GGGATGTGTATCTCGCTCGTCAAGCTGGCTCGGCTGCACATGATGAGACACCGCG 87994
QY 1218 GAGCCAGATACCTTGGGCGATCGGCTAAGCATGCGCTCCACCTCGCTACGTCGCG 1277
Db 87993 GAGGCGAGCGAGCGGCTGTGGGTCTGAGCATCGCATGCTGTGGTTCGTGGCG 87934
QY 1278 TTCTTCTCCATCGCTTGGCCCATCACCTGGGTGTACAGCTCGGAGATCTTCCGCTC 1337
Db 87933 TCCTTCTCCATCGGATGGCGCGGATCGGTGTACAGCTCGGAGATCTTCCGCTG 87874
QY 1338 CAGTGGCGCGCTGGCTCTGCTGGGTGTGGCGTGCAGCGCTACAGCGCGCTC 1397
Db 87873 CGCTGGCGCGAGGATCGGCTCGGACAGCGATGAACCGGTGTGGTGAAGCGCGC 87814
QY 1398 ATCTCATGACCTTCTCTGCTGTCTCAAGCCATCACCATCGGCGGAGCTTCTCTC 1457
Db 87813 GTAGCATGTCTTCTCTGCTGTCTCAAGCCATCACCTTCCGCGGAGCTTCTCTC 87754
QY 1458 TACTCGGCATCGCGCGCTCGCTGGGTCTTCTTACACCTACCTCGGAGACCGCG 1517
Db 87753 TAGCGCGCATAGCTCGCGCGGTGGGTCTCATGTTCTTCTCTCTCGCGGAGCGCA 87694
QY 1518 GCGCGAGCTGTGAGGATGACAGACTCTTCCGCGAC 1556
Db 87693 GGAGGAGCTGTGAGGACACCGTCAAACTCTTCCGCGCG 87655
```

```
RESULT 15
AC115687/c 181163 bp DNA linear PLN 22-MAR-2002
LOCUS
DEFINITION Genomic sequence for Oryza sativa, Nipponbare strain, clone
OSJNBa0091J11, from chromosome 3, complete sequence.
```

```
AC115687
VERSION AC115687.1 GI:19683096
KEYWORDS HTG.
```

```
SOURCE
ORGYA sativa (japonica cultivar-group)
ORGYA sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
```

```
REFERENCE
1 (bases 1 to 181163)
McCombie,W.R., de la Bastide,M., Spiegel,L., Preston,R.,
Ferraro,K., Kuit,K., Nascimento,L., Zlatavarn,T., Balija,V.,
Bell,M., Baker,J., Santos,L., Miller,B., Katzenberger,F.,
Muller,S., King,L., Yang,C., Dike,S., O'Shaughnessy,A., Palmer,L.
and Bedhia,N.
```

```
TITLE Genomic sequence for Oryza sativa, Nipponbare strain, clone
OSJNBa0091J11, from chromosome 3, complete sequence
JOURNAL Unpublished
2 (bases 1 to 181163)
```

```
REFERENCE
McCombie,W.R.
Direct Submission
Submitted (22-MAR-2002) Lita Annenberg Hazen Genome Center, Cold
Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
NY 11724, USA
```

```
COMMENT This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.
```

```
FEATURES
Location/Qualifiers
1..181163
/or-ganism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBa0091J11"
/clone_1lb="HindIII"
```

```
misc_feature 52283..52374
/note="We believe the assembly to be correct. The
sequence is a dinucleotide (TA) repeat followed by a
dinucleotide (CA) repeat. Four subclones span into unique
sequence on both sides."
53504..53580
/note="We believe the assembly to be correct. The
sequence is a dinucleotide (TA) repeat followed by a
dinucleotide (CA) repeat. Three subclones span into
unique sequence on both sides."
132718..132738
/note="We believe the assembly to be correct. The
sequence is a mononucleotide (C) repeat with a single
unique base (A) followed by another mononucleotide (C)
repeat. A single subclone spans the region into unique
sequence on both sides."
152459..152479
/note="We believe the assembly to be correct. The
sequence is a mononucleotide (C) repeat flanking a single
unique base (A) which is followed by an additional
mononucleotide (C) repeat. A single subclone spans
through to unique sequence on both sides."
```

ORIGIN

```
Query Match 37.6%; Score 759.4; DB 8; Length 181163;
Best Local Similarity 73.0%; Pred. No. 4.8e-94;
Matches 992; Conservative 1; Mismatches 357; Indels 9; Gaps 1;

QY 207 GATATCGGGGTGATGAGCGGGCGTCTGCTGTACATCAAGAAGGAGCTTCAACATCATGTGAC 266
Db 175826 GACATCTCGGTGATGAGCGGCGGAGAGTTTCAAGAGAGGACCTGAACATCTCGGAC 175767
QY 267 GGAAGGTGGAGTTCTCATGGGCATCTGACCTTCTACTGCTCATGGCTCTCTCGG 326
Db 175766 GGAAGGTGGAGTTCTCTCGCGGCATCAACATCTACTGCTGCTGGTCTCGG 175707
QY 327 GCGGGCGGAGCTCGGACTGATCGCGCGGGTACACCATCTGTCTTCGCGCGCTCAT 386
Db 175706 GCGGCGGAGCTCGGACTGATCGGCGTCTGCTACACCATGCTGCTCGGTGCTCAT 175647
QY 387 TTCTTCGGGGGGTTCCTCATGGGTTTCGGTTCAGCTGCTGCTGCTGCTGCTGCTGCTG 446
Db 175646 TTCTTCGGGGGGTTCCTCATGCGGCTCATGCGGCTCGCGGAGTACGCGATCGTATGCTCGG 175587
QY 447 CGCTTCGGGGCGGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 506
Db 175586 CGCTTCGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175527
QY 507 GAGGTGTCGCGCGCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566
Db 175526 GAGGTGTCGCGCGCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175467
QY 567 TTCGCGATCTGCTCGGTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
Db 175466 GCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175407
QY 627 GGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 686
Db 175406 GGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175347
QY 687 CTCGCGATCGCGGAGTTCGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
Db 175346 CTCGCGATCGCGGAGTTCGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175287
QY 747 GTGGTGTGGAGAGACCTCGACACCGGAGAGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTG 806
Db 175286 CGTGTCTCGAGAGACCTCGACCTCGGCGGCGGAGGCGGCTGCTGCTGCTGCTGCTGCTG 175227
QY 807 AAGGCGCGCGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857
Db 175226 AAGACGCGGTGCGCATCCCCGAGGAGTCTCCGACGAGGAGCGAAGTTGTAGCGCTGCTG 175167
QY 858 CCCAAGAGAGGAGCGGAAACGAGNAGCGGGTGTGGAGGAGCTCATCTGTCCCGGAC 917
```

Db	175166	CACAAAGAGAGGGGCTCCACCGGAGGGGTCTGGAGGACCTGTGCTCCCGCCGAGC	175107
Qy	918	CCGCCCATGCGCGCATCTCTCTCCGGGATCGGCATCCACTTCTTCCAGCATGGGTTG	977
Db	175106	CCGCCGCTGCGCGCATCTCTCTCCGGGATCGGCATCCACTTCTTCCAGCATGGGTTG	175047
Qy	978	GGCATTCACCTCCGCTCTCTACAGCCCTCTCTGTTCAAGAGCCCGGATTAAAGAAC	1037
Db	175046	GGCATTCACCTCCGCTCTCTACAGCCCTCTCTGTTCAAGAGCCCGGATTAAAGAAC	174987
Qy	1038	GACAAACACATTTCTGGGACACCATTTGGCGTTCGGTCTCACCAAGAGGCTTTTCACTTTG	1097
Db	174986	GACTCCGACTCCATCCGAGGCTCGTGCGGTGGGCGCCAGCAAGACGCTCTTCACTCTC	174927
Qy	1098	TTGGCGACTTCTTCATCGAGCGGCTGGGCGGCGCGGCTGTGCTGGGACACAGGGC	1157
Db	174926	GTGGCCACGTTCTCTCTCGACCGGCTCGGCGGAGGCGGCTGTCTCTCAACAGCGCGGC	174867
Qy	1158	GGGATAATCTCTCTCTCTCATCGGCTCGGCGCGGCTCACCGTCTGTCGGCCACACCCC	1217
Db	174866	GGGATGGTGTCTCTCTCTCATCGGCTCGGCGCGGCTCACCGTCTGTCGGCCACACCCC	174807
Qy	1218	GACCCCAAGATACCTTGGGCGCATCGGCTTAAGCATCGCCCTCCACCTCGGCTAGTCGCC	1277
Db	174806	GAGGCCAGGCGACGGGCGCTGGTGGTCTGAGCATCGCGATGGTGTGTCTGTGGCG	174747
Qy	1278	TTCTTCTCCATCGGCTTGGCCCATCACGTTGGGTGTACAGCTCGGAGATCTTCCCGCTC	1337
Db	174746	TCTTCTCCATCGGATGGGCGCGATCGGTTGGGTGTACAGCTCGGAGATCTTCCCGCTG	174687
Qy	1338	CAGGTGCGCGCTGGGTGTCTCTCGGCTCGGCGCGGCTAACCGGTCACACAGCGGCTC	1397
Db	174686	CGGCTGCGCGCGAGGATGGCGCTCGGACCGCGATGAACCGGTCGTGAGCGCGCC	174627
Qy	1398	ATCTCCATGACTTCTCTCTCTCTCAAGCCCATCACCATCGGCGGAGCTTCTTCTC	1457
Db	174626	GTCAGCATGTCTTCTCTCTCTCTCAAGCCCATCACCATCGGCGGAGCTTCTTCTC	174567
Qy	1458	TACTCCGGCATCGCGCGCTCGCTGGGTGTCTTCTACACCTACCTCCCGGAGACCCGC	1517
Db	174566	TACCGCGCATAGCTCGGCGCGGTGGGTGTTCATGTTCTTCTCTGCGGAGACGCAA	174507
Qy	1518	GGCGGAGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC	1556
Db	174506	GGCAGGAGCTGGAGGACACCGTCAAACTCTTCGGCGGC	174468

Search completed: July 2, 2004, 06:17:20
Job time : 12609 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 00:48:19 ; Search time 1156 Seconds
(without alignments)
7412.297 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 2017
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2016.6	100.0	2017	6	Abk51972 Rice cont
2	2016.6	100.0	2017	7	Abx93208 CDNA enco
3	1094.6	54.3	2089	6	Abk51974 Wheat cdn
4	1094.6	54.3	2089	7	Abx93210 CDNA enco
5	1005	49.8	1914	6	Abk51971 Corn CDNA
6	1005	49.8	1914	7	Abx93207 CDNA enco
7	981.6	48.7	1872	6	Abk51975 Wheat cdn
8	981.6	48.7	1872	7	Abx93211 CDNA enco
9	575.2	28.5	1853	6	Abk51973 Soybean c
10	575.2	28.5	1853	7	Abx93209 CDNA enco
11	477.2	23.7	2056	6	Abk51974 Nucleotid
12	470.2	23.3	1527	3	Aac43261 Arabidops
13	469.2	23.3	1925	6	Abk51975 Nucleotid
14	443.2	22.0	1662	6	Abk51976 Nucleotid
15	433.8	21.5	1482	6	Abk51990 Arabidops
16	426	21.1	1766	6	Abk51977 Nucleotid
17	397.4	19.7	1715	6	Abk51978 Nucleotid
18	387.6	19.2	1722	6	Abk51979 Nucleotid
19	343.6	17.0	615	7	Abk51970 Corn CDNA
20	343.6	17.0	615	7	Abx93206 CDNA enco
21	325.4	16.1	1644	6	Abk51971 Arabidops
22	325	16.1	1690	6	Abk51972 Nucleotid
23	313.2	15.5	756	8	Ada49209 Maize gen

24	295	14.6	1004	8	ADA48319	Rice gene
25	191.2	9.5	2000	7	ADA72262	Rice gene
26	188.8	9.4	1675	2	AAZ32195	Corn hexo
27	172.4	8.5	1560	7	ADA70449	Rice gene
28	170	8.4	1377	7	ADA70153	Rice gene
29	168.4	8.3	1776	2	AAZ32200	Rice hexo
30	164.2	8.1	616	8	ADA49126	Wheat gen
31	159.8	7.9	349980	6	ABQ81849	Abq81849 Bifidobac
32	158.2	7.8	1539	7	ADA70241	Rice gene
33	152.8	7.6	1350	7	ACA37723	Prokaryot
34	148.8	7.4	1545	7	ACA40941	Prokaryot
35	147.2	7.3	1509	7	ACA40810	Prokaryot
36	147.2	7.3	110000	4	AAI99682_37	Continuation (38 o
37	141.2	7.0	1752	2	AAZ32203	Wheat hex
38	139.8	6.9	1572	7	ADA70338	Rice gene
39	136.4	6.8	1577	8	ADA48271	Rice gene
40	136.4	6.8	1577	9	ADC07783	Rice DNA
41	136	6.7	1167	2	AAZ32199	Rice hexo
42	130.2	6.5	1395	7	ACA32312	Prokaryot
43	129.8	6.4	1281	7	ACA38748	Prokaryot
44	128	6.3	110000	4	AAI99683_37	Continuation (38 o
45	121.4	6.0	2504	4	AAF55869	Rat GLUTX

ALIGNMENTS

RESULT 1
ABK51972
ID ABK51972 standard; cDNA; 2017 BP.
XX
AC ABK51972;
XX
DT 27-AUG-2002 (first entry)
XX
DE Rice contig encoding Beta vulgaris-like sugar transport protein.
XX
KW Rice; Beta vulgaris-like sugar transport protein; carbohydrate transport;
XX grain filling; annual field crop; plant; gene; ss.
XX Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 96..1628
FT /*tag= a
FT /product= "Rice Beta vulgaris-like sugar transport
FT protein"
XX
XX US6393776-B1.
XX
XX PD 07-MAY-2002.
XX
XX PF 14-APR-1999; 99US-00291922.
XX
XX PR 24-APR-1998; 98US-0083044P.
XX
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX DR WPI; 2002-453364/48.
XX
XX DR P-P8DB; AAU97211.
XX
XX PT New nucleic acid encoding plant sugar-transport proteins, useful for
XX preparing transgenic plants with altered carbohydrate distribution.
XX
XX PS Example 4; Col 61-64; 54pp; English.
XX
XX CC The present invention relates to the isolation of plant polynucleotide
XX sequences encoding an Arabidopsis thaliana-like sugar transport protein
XX or Beta vulgaris-like sugar transport protein. The polynucleotide
XX sequences are useful for altering the level of sugar transport proteins
XX in plants, i.e. for control of carbohydrate transport and distribution in


```
Db 1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
RESULT 2
ABX93208
ID ABX93208 standard; cDNA; 2017 BP.
XX AC ABX93208;
XX 29-MAY-2003 (first entry)
XX cDNA encoding rice sugar transport protein #3.
XX
XX Beta vulgaris-like sugar transport protein; corn; rice; wheat;
XX plant sugar transport protein; carbohydrate transport; soybean;
XX carbohydrate distribution; plant; gene; ss.
XX
XX Oryza sativa.
XX
XX US2002178468-A1.
XX
XX 28-NOV-2002.
XX
XX 17-JAN-2002; 2002US-00051902.
XX
XX 24-APR-1998; 98US-0083044P.
XX 14-APR-1999; 99US-00291922.
XX
XX (ALLE/) ALLEN S M.
XX (HITZ/) HITZ W D.
XX (KINS/) KINNEY A J.
XX (TING/) TINGEY S V.
XX
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2003-340957/32.
XX P-PSDB; ABU08336.
XX
XX Novel plant sugar transport proteins and nucleic acid encoding the
XX protein useful for producing transgenic plants having altered levels of
XX sugar transport protein.
XX
XX Claim 7; Page 32-33; 56pp; English.
XX
XX The present invention relates to the isolation of Arabidopsis thaliana-
XX like or Beta vulgaris-like sugar transport proteins, and the
XX polynucleotide sequences encoding them. The plant sugar transport
XX proteins of the invention have been isolated from corn, rice, soybean,
XX and wheat. The polypeptides of the invention may be used for altering the
XX level of expression of a sugar transport protein in a host cell, by
XX transforming a host cell with a chimeric construct encoding all, or a
XX portion of the sugar transport protein, in sense or antisense
XX orientation. Particularly, the polypeptides may provide a means to
XX control carbohydrate transport and distribution in plants. ABX93206-
XX ABX93211 represent cDNA sequences encoding Beta vulgaris-like sugar
XX transport proteins
XX
XX Sequence 2017 BP; 389 A; 527 C; 582 G; 418 T; 0 U; 1 Other;
XX
XX Query Match 100.0%; Score 2016.6; DB 7; Length 2017;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 2017; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTACATGTAAGCTCGTGGCGGACGAGCTTACACTCGACGCCACTACTGTACAGGCC 60
Db 1 CTTACATGTAAGCTCGTGGCGGACGAGCTTACACTCGACGCCACTACTGTACAGGCC 60
QY 61 CAGAGCGAGCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 61 CAGAGCGAGCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 121 CCGTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
```

Db 1201 TCGTCGGCAGCACCCGAGCCCAAGATACTTGGGCCATCGGCTTAAGCATCGCCTCCA 1260
 Qy 1261 CCCTCGCTAGTGCCTTCTTCCATCGCCTTGGCCCATCATCATCGTGGGTGACAGCT 1320
 Db 1261 CCCTCGCTAGTGCCTTCTTCCATCGCCTTGGCCCATCATCATCGTGGGTGACAGCT 1320
 Qy 1321 CGGAGATCTTCCGCTCCAGGTGGCGCGCTGGGCTGCTCGCTCGGCGTGGCCCAACC 1380
 Db 1321 CGGAGATCTTCCGCTCCAGGTGGCGCGCTGGGCTGCTCGCTCGGCGTGGCCCAACC 1380
 Qy 1391 CGGTCCAGCGGGGTATCTTCATGACCTTCCTGTGCTGTGTCGAGGCAATCACATCG 1440
 Db 1391 CGGTCCAGCGGGGTATCTTCATGACCTTCCTGTGCTGTGTCGAGGCAATCACATCG 1440
 Qy 1441 CGCGCAGCTTCTCTCTACTCCGGCATCGCGGCTCGGCTGGGTGCTTCTTACACCT 1500
 Db 1441 CGCGCAGCTTCTCTCTACTCCGGCATCGCGGCTCGGCTGGGTGCTTCTTACACCT 1500
 Qy 1501 ACCTCCCGAGACCGCGCGCGGCGAGCTGTGGAGAGATGAGCAAGCTTTCGGGACACGG 1560
 Db 1501 ACCTCCCGAGACCGCGCGCGGCGAGCTGTGGAGAGATGAGCAAGCTTTCGGGACACGG 1560
 Qy 1561 CGCGCCTCGGATCAGCAGCGCCGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 Db 1561 CGCGCCTCGGATCAGCAGCGCCGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 Qy 1621 CTAAGTATCAAACTAACCGCAAAATCACCAAACTCTAAGGGTTTTCTTCAAAAAGCTG 1680
 Db 1621 CTAAGTATCAAACTAACCGCAAAATCACCAAACTCTAAGGGTTTTCTTCAAAAAGCTG 1680
 Qy 1691 TGCTGTACTGCTAGCTAGCAAGTAGTAGCAGCAAGCTGGAGAGATTCGCTGATCGGGG 1740
 Db 1691 TGCTGTACTGCTAGCTAGCAAGTAGTAGCAGCAAGCTGGAGAGATTCGCTGATCGGGG 1740
 Qy 1741 TTGCTGAGAGCGAGCGCGCGGCGAGCAAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800
 Db 1741 TTGCTGAGAGCGAGCGCGCGGCGAGCAAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800
 Qy 1801 TCATCTTCAAGTATGATGATTTATTTTGTCTTTTGTCTTTGTCTTTGTCTTTGTCTAT 1860
 Db 1801 TCATCTTCAAGTATGATGATTTATTTTGTCTTTTGTCTTTGTCTTTGTCTTTGTCTAT 1860
 Qy 1861 GCGATGAAGATACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
 Db 1861 GCGATGAAGATACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
 Qy 1921 GTCAGTGTCTTGTGTAAGAAATGTTTAACTGTTTAACTGTTTAACTGTTTAACTGTT 1980
 Db 1921 GTCAGTGTCTTGTGTAAGAAATGTTTAACTGTTTAACTGTTTAACTGTTTAACTGTT 1980
 Qy 1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
 Db 1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017

RESULT 3

ABK51974

ID ABK51974 standard; cDNA; 2089 BP.

XX AC ABK51974;

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

FT /tag= a
 FT /partial
 FT /product= "wheat Beta vulgaris-like sugar transport
 FT protein"
 FT /note= "This sequence lacks a start codon"
 XX
 PN US6383776-B1.
 XX
 PD 07-MAY-2002.
 XX
 PF 14-APR-1999; 99US-00291922.
 XX
 PR 24-APR-1998; 98US-0083044P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 DR WPI; 2002-453364/48.
 DR P-PSDB; AAU97213.
 XX
 PT New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution.
 XX
 PS Example 4; Col 71-72; 54pp; English.
 XX
 CC The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
 CC or beta vulgaris-like sugar transport protein. The polynucleotide
 CC sequences are useful for altering the level of sugar transport proteins
 CC in plants, i.e. for control of carbohydrate transport and distribution in
 CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
 CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
 CC sugar transport. The polynucleotide sequences can also be used to isolate
 CC cDNA sequences and genes that encode homologues of the new proteins. The
 CC present sequence encodes a wheat Beta vulgaris-like sugar transport
 CC protein
 XX
 SQ Sequence 2089 BP; 395 A; 672 C; 605 G; 417 T; 0 U; 0 Other;
 Query Match 54.3%; Score 1094.6; DB 6; Length 2089;
 Best Local Similarity 79.1%; Pred. No. 1.9e-177;
 Matches 1343; Conservative 1; Mismatches 335; Indels 18; Gaps 3;
 Qy 84 GCACACCGGAGATGGCTTCGCCCGCTCGCGAGGCGCTCGGCCCAAGAGAGGCG 143
 Db 65 GGGTCGTGAGATGAGACCGCGCGCACTCCGCGCGCGCTCGAGCCCAAGAGAGGCG 124
 Qy 144 AACGTCCGGTTCGCTTCGCTCGCCATCTCGCTCCATGACCTCCATCTCTCTCGGC 203
 Db 125 AACGTGAGTTCGCTTCGCTTCGCTCGCCATCTCGCTCCATGACCTCTCTCTCGGC 184
 Qy 204 TACGATATCGGGGTGATGAGCGGGCGTGGTGTATCATCAAGAGAGACTTCAACATAGT 263
 Db 185 TACGATATCGGGGTGATGAGCGGGCGTGGTGTATCATCAAGAGAGACTTCAACATAGT 244
 Qy 264 GACGGAAAGTGGAGGTTCTCATGGGCATCTGAACCTTACTCGCTCATCGGTCCTTC 323
 Db 245 GACACCGAGTGGAGTCTCATGGGCATCTCAACGTGTACTGCTCATTTGGTCTCTTC 304
 Qy 324 GCGCGGGCGGAGCGTGGAGTGGATCGCGGGGTACACATCGTGTTCGCGCGGTC 383
 Db 305 GCGCGGGCGGAGCGTGGAGTGGATCGCGGGGTTCACCATCGTCTTCGCGCGGTC 364
 Qy 384 ATATTCTTCGGGGGSGTTCTCATGGGTTCGCGGTCAACTACGATCTCATGTTTC 443
 Db 365 ATCTTCTTCGGGGGCGCTTCATGAGGCTTCGCGTCAACTACGATCTCATGTTTC 424
 Qy 444 GCGCGCTTCGTCGGCGGATCGCGGTGAGGCTACGCGTTCATGATCGCGCGGTGTACAC 503
 Db 425 GCGCGCTTCGTCGGCGGATCGCGGTGAGGCTACGCTCTCATGATCGCGCGGTGAACAC 484
 Qy 504 GCCGAGGTGTTCGCGCGGCTCGCGGTGGCTTCTGAGCTGTTCCCGGAGGTGTTTCATC 563

Wheat cDNA clone wl8.pk0001.all encoding sugar transport protein.

Wheat; Beta vulgaris-like sugar transport protein;

carbohydrate transport; Grain filling; annual field crop; plant;

clone wl8.pk0001.all; gene; ss.

Triticum aestivum.

OS

XX

XX

XX

XX

XX

XX

XX

[illegible]

Db	1565	CGCGAAGCCAGACGACGAGCGCGGGAGAGAAAGGTGGAATGCGCGCCACCACTGACCG	1621
Qy	1641	CAAAATCACCAAAATCCTAAGGGTTTTTCTTCCAAAAACGTGTGCTGTACTGGCTAGCTAGC	1700
Db	1625	CAAGTTGGCAGATCGCGATCGGAAGACTTCGCGTGTATCCGCTTCGCTAGCTAGCTGCC	1694
Qy	1701	A-----ACTAGTAGCAGCAACGTGGGAAGATTCCGCTGATCCGGC---GTTGCT	1745
Db	1685	ACAAGGCCCATAGATACGAAGTAGCGGTGGGAAGATTCTGATCCGCGGAGCTGCC	1744
Qy	1746	GGAGAGCGACGGCGCGC	1762
Db	1745	GGAGGCGACGGCAAGC	1761
RESULT 4			
ABX93210			
ID	ABX93210 standard; cDNA; 2089 BP.		
XX	AC	ABX93210;	
XX	DT	29-MAY-2003 (first entry)	
XX	DE	cDNA encoding wheat sugar transport protein #4.	
KW	Beta vulgaris-like sugar transport protein; corn; rice; wheat;		
KW	plant sugar transport protein; carbohydrate transport; soybean;		
KW	carbohydrate distribution; plant; gene; ss.		
OS	Triticum aestivum.		
XX	PN	US2002178468-A1.	
XX	PD	28-NOV-2002.	
XX	XX	17-JAN-2002; 2002US-00051902.	
PF			
XX	XX	24-APR-1998; 98US-0083044P.	
PR	PR	14-APR-1999; 99US-00291922.	
XX	XX	(ALLE/) ALLEN S M.	
PA	PA	(HITZ/) HITZ W D.	
PA	PA	(KINN/) KINNEY A J.	
PA	PA	(TING/) TINGEY S V.	
XX	PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV;	
XX	XX	WPI: 2003-340957/32.	
DR	DR	P-PSDB; ABU08338.	
XX	Novel plant sugar transport proteins and nucleic acid encoding the		
PT	PT	protein useful for producing transgenic plants having altered levels of	
PT	PT	sugar transport protein.	
XX	XX	Claim 7; Page 37-38; 56pp; English.	
PS	The present invention relates to the isolation of Arabidopsis thaliana-		
XX	CC	like or Beta vulgaris-like sugar transport proteins, and the	
CC	CC	polynucleotide sequences encoding them. The plant sugar transport	
CC	CC	proteins of the invention have been isolated from corn, rice, soybean,	
CC	CC	and wheat. The polypeptides of the invention may be used for altering the	
CC	CC	level of expression of a sugar transport protein in a host cell, by	
CC	CC	transforming a host cell with a chimeric construct encoding all, or a	
CC	CC	portion of the sugar transport protein, in sense or antisense	
CC	CC	orientation. Particularly, the polypeptides may provide a means to	
CC	CC	control carbohydrate transport and distribution in plants. ABX93206-	
CC	CC	ABX93211 represent cDNA sequences encoding Beta vulgaris-like sugar	
CC	CC	transport proteins	
XX	XX	Sequence 2089 BP; 395 A; 672 C; 605 G; 417 T; 0 U; 0 Other;	
SQ			
Query Match 54.3%; Score 1094.6; DB 7; Length 2089;			
Best Local Similarity 79.1%; Pred. No. 1.9e-177;			

Query Match 54.3%; Score 1094.6; DB 7; Length 2089;
Best Local Similarity 79.1%; Pred. No. 1.9e-177;

Matches 1343; Conservative 1; Mismatches 335; Indels 18; Gaps 3;

```

QY      84  GCACACCGAGATGGCTTCGCGCGGCTGCGGAGCGGTGCGCGGAGAGAGAGGGC 143
DB      84  GCACACCGAGATGGCTTCGCGCGGCTGCGGAGCGGTGCGCGGAGAGAGAGGGC 143
QY      65  GCGTGTGAGGATGGAGCGCGCGGCACTCCGCGCGCGCGTTCGAGCGCAAGAGAGGGC 124
DB      65  GCGTGTGAGGATGGAGCGCGCGGCACTCCGCGCGCGCGTTCGAGCGCAAGAGAGGGC 124
QY      144 AACGTCCGGTTCGCTTCGCGCGGCGCATCTCGCGTCCATGACCTCCATCTCTCTCGGC 203
DB      144 AACGTCCGGTTCGCTTCGCGCGGCGCATCTCGCGTCCATGACCTCCATCTCTCTCGGC 203
QY      125 AACGTGAGGTTGCTTCGCGCGGCGCATCTCGCGTCCATGACCTCCATCTCTCTCGGC 184
DB      125 AACGTGAGGTTGCTTCGCGCGGCGCATCTCGCGTCCATGACCTCCATCTCTCTCGGC 184
QY      204 TACGATATCGGGGTGATGAGCGGGCGGTGCTGTATCATCAAGAGAGACTTCAACATCAGT 263
DB      204 TACGATATCGGGGTGATGAGCGGGCGGTGCTGTATCATCAAGAGAGACTTCAACATCAGT 263
QY      185 TACGACATCGCGGTGATGAGCGGAGCGTTCGCTGTATCATCAAGAGAGACTTCAAGATCAAC 244
DB      185 TACGACATCGCGGTGATGAGCGGAGCGTTCGCTGTATCATCAAGAGAGACTTCAAGATCAAC 244
QY      264 GACGGGAGGTGGAGGTTCTCATGGGATAGTGAACCTCTACTCTGCTCATCGGCTCTTTC 323
DB      264 GACGGGAGGTGGAGGTTCTCATGGGATAGTGAACCTCTACTCTGCTCATCGGCTCTTTC 323
QY      245 GACACCCAGCTGGAGGTTCTCATGGGATCTCTCAACGTGTACTCTGCTCATTTGGCTCTTTC 304
DB      245 GACACCCAGCTGGAGGTTCTCATGGGATCTCTCAACGTGTACTCTGCTCATTTGGCTCTTTC 304
QY      324 GCGGCGGGCGGAGCTTCGAGTGGATCGGCGGCGGTACACCATCGTGTGTTTCGCGCGGCTC 383
DB      324 GCGGCGGGCGGAGCTTCGAGTGGATCGGCGGCGGTACACCATCGTGTGTTTCGCGCGGCTC 383
QY      305 GCGGCGGGCGGAGCTTCGAGTGGATCGGCGGCGGTTCACCATCGTGTGTTTCGCGCGGCTC 364
DB      305 GCGGCGGGCGGAGCTTCGAGTGGATCGGCGGCGGTTCACCATCGTGTGTTTCGCGCGGCTC 364
QY      384 ATATTCTTCGCGGGGSGTTCCTCATGGGTTTCGCGGTCAACTAGCCATGCTCATGTTTC 443
DB      384 ATATTCTTCGCGGGGSGTTCCTCATGGGTTTCGCGGTCAACTAGCCATGCTCATGTTTC 443
QY      365 ATCTTCTTCGCGGGCGGCTTCATCATGGGCTTCTCCGTCAACTAGCCATGCTCATGTTTC 424
DB      365 ATCTTCTTCGCGGGCGGCTTCATCATGGGCTTCTCCGTCAACTAGCCATGCTCATGTTTC 424
QY      444 GCGCGTTCGTCGCGGCGGATCGGCGTGGGTACGCGTCTATGATCGCGCGGCTGTACACC 503
DB      444 GCGCGTTCGTCGCGGCGGATCGGCGTGGGTACGCGTCTATGATCGCGCGGCTGTACACC 503
QY      425 GGGCGTTCGTCGCGGCGGATCGGCGTGGGTACGCGTCTATGATCGCGCGGCTGTACACC 484
DB      425 GGGCGTTCGTCGCGGCGGATCGGCGTGGGTACGCGTCTATGATCGCGCGGCTGTACACC 484
QY      504 GCCAGGTGTCGCCGCGGTGCGCGGTGCTTCGTGACGTGTTTCGCGGAGGTGTTTCATC 563
DB      504 GCCAGGTGTCGCCGCGGTGCGCGGTGCTTCGTGACGTGTTTCGCGGAGGTGTTTCATC 563
QY      485 GCGGAGGTGTCGCCGCGGTGCGCGGTGCTTCGTGACGTGTTTCGCGGAGGTGTTTCATC 544
DB      485 GCGGAGGTGTCGCCGCGGTGCGCGGTGCTTCGTGACGTGTTTCGCGGAGGTGTTTCATC 544
QY      564 AACTTCGCGCATCTGCTCGGATGCTCGAAGTATGCTTCTCCCGTTCGCGGTGAAAC 623
DB      564 AACTTCGCGCATCTGCTCGGATGCTCGAAGTATGCTTCTCCCGTTCGCGGTGAAAC 623
QY      545 AACTTCGCGCATCTGCTCGGATGCTCGAAGTATGCTTCTCCCGTTCGCGGTGAAAC 604
DB      545 AACTTCGCGCATCTGCTCGGATGCTCGAAGTATGCTTCTCCCGTTCGCGGTGAAAC 604
QY      624 CTCGGTGGCGCATCATGCTCGGATCGGCGGCGGCGGCTCGGTGCTGCTCGCGCTCATG 683
DB      624 CTCGGTGGCGCATCATGCTCGGATCGGCGGCGGCGGCTCGGTGCTGCTCGCGCTCATG 683
QY      605 CTCGGTGGCGCATCATGCTCGGATCGGCGGCGGCGGCTCGGTGCTGCTCGCGCTCATG 664
DB      605 CTCGGTGGCGCATCATGCTCGGATCGGCGGCGGCGGCTCGGTGCTGCTCGCGCTCATG 664
QY      684 GTGCTCGGATCGCGGATGCTCGCGGCTGCTGCTGATGAGGACGCTTCGCGGAGCGC 743
DB      684 GTGCTCGGATCGCGGATGCTCGCGGCTGCTGCTGATGAGGACGCTTCGCGGAGCGC 743
QY      665 GTGCTCGGATCGCGGATGCTCGCGGCTGCTGCTGATGAGGAGCGCTTCGCGGAGCGC 724
DB      665 GTGCTCGGATCGCGGATGCTCGCGGCTGCTGCTGATGAGGAGCGCTTCGCGGAGCGC 724
QY      744 AAGGTGTGCTGGAGAGACCTCCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGAGC 803
DB      744 AAGGTGTGCTGGAGAGACCTCCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGAGC 803
QY      725 AAGGTGTGCTGGAGAGACCTCCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGAGC 784
DB      725 AAGGTGTGCTGGAGAGACCTCCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGAGC 784
QY      804 ATCAAGGCGCGCGCGGATCCCTGAGAGCTCGAGCGGAGCGGTGAGACCGTCCCGC --- 860
DB      804 ATCAAGGCGCGCGCGGATCCCTGAGAGCTCGAGCGGAGCGGTGAGACCGTCCCGC --- 860
QY      785 ATTAAGACTGCGCGCGGATCCCTCTGGGCTTCGACGCGGAGCGGTGCTCCCGTCCGCA 844
DB      785 ATTAAGACTGCGCGCGGATCCCTCTGGGCTTCGACGCGGAGCGGTGCTCCCGTCCGCA 844
QY      861 AAGAGAGGAGCGGAAACGAGAGCGGTGTGGAAGGAGTCTATCTGTCGCCGAGCGCGC 920
DB      861 AAGAGAGGAGCGGAAACGAGAGCGGTGTGGAAGGAGTCTATCTGTCGCCGAGCGCGC 920
QY      845 AACAAAGGAGCAGCAGGAGGAGCGGTTTGAAGAGCTCTATCTGTCACCGAGCATTA 904
DB      845 AACAAAGGAGCAGCAGGAGGAGCGGTTTGAAGAGCTCTATCTGTCACCGAGCATTA 904
QY      921 GCCATGCGGCGCATCTCTGCTGTGCGGATCGGATCCGATCCACTTCTTCCAGCATGCGTTGGC 980
DB      921 GCCATGCGGCGCATCTCTGCTGTGCGGATCGGATCCGATCCACTTCTTCCAGCATGCGTTGGC 980
QY      905 GCCATGCGGCGCATCTCTGCTGTGCGGATCGGATCCGATCCACTTCTTCCAGCATGCGTTGGC 964
DB      905 GCCATGCGGCGCATCTCTGCTGTGCGGATCGGATCCGATCCACTTCTTCCAGCATGCGTTGGC 964
QY      981 ATTTCATCTCGCTGCTTCTACGCGCTCTGCTGTGTCAGAGCGCGGATTAAGAGACGAC 1040
DB      981 ATTTCATCTCGCTGCTTCTACGCGCTCTGCTGTGTCAGAGCGCGGATTAAGAGACGAC 1040
QY      965 ATCAGCGCGGCTGCTGCTCTACGCGCGGCTAGTTTTCAGAGCGCGGATTAAGAGACGAC 1024
DB      965 ATCAGCGCGGCTGCTGCTCTACGCGCGGCTAGTTTTCAGAGCGCGGATTAAGAGACGAC 1024
QY      1041 AAACATCTTCTGGCAGCACCATTGCGCGCTGCTGCTGTCACCAAGAGGCTTTTTCATCTTTGTTG 1100
DB      1041 AAACATCTTCTGGCAGCACCATTGCGCGCTGCTGCTGTCACCAAGAGGCTTTTTCATCTTTGTTG 1100
QY      1025 AGCGCTCTCGCGCGCACCGTGGCGGTTCGGGCGCACCAATACGCTCTTTCATCTCTGGTG 1084
DB      1025 AGCGCTCTCGCGCGCACCGTGGCGGTTCGGGCGCACCAATACGCTCTTTCATCTCTGGTG 1084
QY      1101 GCGACTTCTTCATCGAGGCGGTTCGGGCGGCGGCTGCTGCTGCGGAGCAGCGGCGGG 1160
DB      1101 GCGACTTCTTCATCGAGGCGGTTCGGGCGGCGGCTGCTGCTGCGGAGCAGCGGCGGG 1160
QY      1085 GCGACCTTCTCTCTCGAGCGGATCCGCGCGGCGGCGGCTGCTGCTGACGACGCGGCGGC 1144
DB      1085 GCGACCTTCTCTCTCGAGCGGATCCGCGCGGCGGCGGCTGCTGCTGACGACGCGGCGGC 1144

```

```

QY      1161 ATAAATCTCTCTCTCATCGGCTTCGCGCGGCTTCACGCTCGTTCGCGGAGCAGCAGCGGAC 1220
DB      1145 ATGCTCTGCTCTCTTATGTTGGCTTCGCGAGCGGGGCTCACGCTCATCAGCGCGCAGCGGAC 1204
QY      1221 GCAAAGATACCTTTGGGCCATCGGCTTAAGCATCGGCTTCACGCTTCGCTTCAGTTCGCTTC 1280
DB      1205 GAGAGATACCTTTGGGCCATCGTCTGTGCTGCTTCGATCATGCTTCAGTTCGCTTC 1264
QY      1281 TTCTCATTCGCTTCGCGCCCATACGTTGGGTGTAAGTTCGAGATCTTCCGCTTCAG 1340
DB      1265 TTCTCATTCGCTTCGCGCCCATACGTTGGGTGTAAGTTCGAGATCTTCCGCTTCAG 1324
QY      1341 GTGCGCGCTCGGCTGCTCGCTCGGCTTCGCGGCAACCGGCTCACGAGCGGCTCATC 1400
DB      1325 GTGCGCGCTCGGCTGCTCGCTCGGCTTCGCGGCAACCGGCTCACGAGCGGCTCATC 1384
QY      1401 TCATGACCTTCTCTGCTGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGCTTC 1460
DB      1385 TCATGACCTTCTCTGCTGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGCTTCGCTTC 1444
QY      1461 TCGGCTATCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 1520
DB      1445 GCGGCTATCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 1504
QY      1521 CGGAGCTTCGAGGATGAGCAAGCTTTCGCGGCAACCGGCTTCGAGGCTTCGAGGCTTCGAG 1580
DB      1505 CGGAGCTTCGAGGATGAGCAAGCTTTCGCGGCAACCGGCTTCGAGGCTTCGAGGCTTCGAG 1564
QY      1581 GAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1640
DB      1565 GCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1624
QY      1641 CAAATACCAATCTTAAGGTTTCTTCAAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
DB      1625 CAAATACCAATCTTAAGGTTTCTTCAAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1684
QY      1701 A-----AGTAGTAGCAGCAACCTGAGGAGATTCGCTGATTCGCTGCTGCTGCTGCTGCTGCTGCT 1745
DB      1685 ACAAGGCGCACATAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1744
QY      1746 GGAGAGCGAGCGCGCGC 1762
DB      1745 GGAGAGCGAGCGCGCAAGC 1761

```

RESULT 5
ABK51971
ID ABK51971 standard; cDNA; 1914 BP.
XX ABK51971;
XX
XX 27-AUG-2002 (first entry)
XX
DE Corn cDNA clone cepe7.pk0018.g3 encoding sugar transport protein.
XX
XX Corn; Beta vulgaris-like sugar transport protein; carbohydrate transport;
KW grain filling; annual field crop; plant; clone cepe7.pk0018.g3; gene; ss.
XX
XX Zea mays.
XX
XX Key Location/Qualifiers
XX 56..1597 /tag= a
XX /product= "Corn Beta vulgaris-like sugar transport protein"
XX
XX US6383776-B1.
XX
XX 07-MAY-2002.
XX
XX 14-APR-1999; 99US-00291922.
XX

Qy	1532	GGAGATAGCAAGCTGTGTCGC-----GACAGCGCCCGCCCTCGGAATCAGACGAGCC	1589
Db	1495	GGAGATGGGCAAGCTGTGTCGCATGCCAGACACGGGCATGCGTGAAGAGCAGAAGAGCC	1554
Qy	1586	AGCCAAGGAGAAAGAAAGGTGGAATGGCGGCCACTAACT	1626
Db	1555	CGAGCCNAGGAGAAGGTGTGGAACTGCTAGCAGCAAGT	1595
RESULT 7			
ABKS1975	ID	ABKS1975 standard; cDNA; 1872 BP.	
XX	ABKS1975;		
DT	27-AUG-2002	(first entry)	
XX	DE	Wheat cDNA clone wlm1.pk0012.h1 encoding sugar transport protein.	
XX	XX	Wheat; Beta vulgaris-like sugar transport protein;	
KW	KW	carbohydrate transport; grain filling; annual field crop; plant;	
KW	KW	clone wlm1.pk0012.h1; gene; ss.	
XX	XX		
OS	OS	Triticum aestivum.	
XX	XX		
Key	Key	Location/Qualifiers	
FF	FF	65..1654	
CDS	CDS	/*tag= a	
FT	FT	/product= "Wheat Beta vulgaris-like sugar transport	
FT	FT	protein"	
FT	FT		
XX	XX		
PN	PN	US6383776-B1.	
XX	XX		
PD	PD	07-MAY-2002.	
XX	XX		
PF	PF	14-APR-1999; 99US-00291922.	
XX	XX		
PR	PR	24-APR-1998; 98US-0083044P.	
XX	XX	(DUPO) DU PONT DE NEMOURS & CO E I.	
PA	PA	Allen SM, Hitz WD, Kinney AJ, Tingey SV;	
PI	PI	WPI; 2002-453364/48.	
DR	DR	P-PSDB; AAU37214.	
XX	XX		
XX	XX	New nucleic acid encoding plant sugar-transport proteins, useful for	
PT	PT	preparing transgenic plants with altered carbohydrate distribution.	
XX	XX	Example 4; Col 75-78; 54pp; English.	
XX	XX		
CC	CC	The present invention relates to the isolation of plant polynucleotide	
CC	CC	sequences encoding an Arabidopsis thaliana-like sugar transport protein	
CC	CC	or Beta vulgaris-like sugar transport protein. The polynucleotide	
CC	CC	sequences are useful for altering the level of sugar transport proteins	
CC	CC	in plants, i.e. for control of carbohydrate transport and distribution in	
CC	CC	plant cells, e.g. during grain filling of annual field crops (e.g. corn,	
CC	CC	rice, soybeans, and wheat), and, for studying carbohydrate flows and	
CC	CC	sugar transport. The polynucleotide sequences can also be used to isolate	
CC	CC	cDNA sequences and genes that encode homologues of the new proteins. The	
CC	CC	present sequence encodes a wheat Beta vulgaris-like sugar transport	
XX	XX	protein	
XX	XX		
SQ	SQ	Sequence 1872 BP; 390 A; 579 C; 534 G; 369 T; 0 U; 0 Other;	
Query Match 48.7%; Score 981.6; DB 6; Length 1872;			
Best Local Similarity 77.4%; Pred No. 3.4e-158;			
Matches 1186; Conservative 1; Mismatches 345; Indels 0; Gaps 0;			
Qy	99	GCTTCGCCCGCTGCGGAGCGCGTCCGCGCGAAGAGGCAAGGCAACGTCGCGTTCGCC	158
Db	128	GCTGCGCTCCGAGCGCGGGGCGAGTCCATCCCAAGCAACGAGCAATTTCAAGTAGCC	187

QY 1239 ATCGGCCTAAGCATCGCCTCCACCTCGCTACGCTTCTTCTTCATCGGCTTGGC 1238
DB 1268 ATCGGCCTGTGCATCGTGTCCATCTTGGCTTACGCTTCTTCTTCCATCGGCTGGG 1327
QY 1299 CCCATCAGTGGGTGATAGCTCGAGATCTTCCGGCTCCAGGTGCGGCGCTGGCTGC 1358
DB 1328 CCCCTCACCGAGGTGTACCTCGAGGTCTTCCACTGCGGGTGCGGCGCTGGCTTC 1387
QY 1359 TCGCTCGGCGTGCCTCGCAACCGCTCACCGCGCTGCTCACTGATGACCTTCTGTCG 1418
DB 1398 GGCCTGGGACGTCATGCAACCGCTCACAGCGCGGCTCTCAATGCTTCTGTCG 1447
QY 1419 CTGTCACAGGCATCACCATCGGCGGAGCTTCTTCTTCTTCTGATCGGCGCTC 1478
DB 1448 TTGTCACAGGCATCACCATCGGCGGAGCTTCTTCTTCTGATCGGCGGATCGGCGGATA 1507
QY 1479 GCCTGGGTGTTCTTCTACACCTTACCTCCCGGAGACCGCGCGGAGATG 1538
DB 1508 GGATGGATTTCTTCTTCACTTCACTCCGAGAGCGCTGCGCTCGAGAGATA 1567
QY 1539 AGCAAGCTGTTCGCGACACCGCGCGCTCGGAATCAGACGAGCCAGCAAGAGAAG 1598
DB 1568 GGAAGCTTTTCGCGATGACGACACCGCGCGCTCGAAGCCCAAGACACCGCCACGAAAGAC 1627
QY 1599 AAGAAGTGGAAATGCGCGCCCACTAATGATCAA 1632
DB 1628 AAGCGAAAGTAGGGAGATGAATAGTGAGCTA 1661

RESULT 8

ID ABX93211 standard; cDNA; 1872 BP.

AC ABX93211;

DT 29-MAY-2003 (first entry)

DE cDNA encoding wheat sugar transport protein #5.

XX Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.

OS Triticum aestivum.

PN US2002178468-A1.

XX 28-NOV-2002.

PF 17-JAN-2002; 2002US-00051902.

PR 24-APR-1998; 98US-0083044P.

PR 14-APR-1999; 99US-00291922.

XX (ALLE/) ALLEN S M.

PA (HITZ/) HITZ W D.

PA (KINN/) KINNEY A J.

PA (TING/) TINGEY S V.

PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

XX WPI; 2003-340957/32.

DR P-PSDB; ABU08339.

XX Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.

PS Claim 7; Page 39-40; 56pp; English.

XX The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport

CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABX93206-
CC ABX93211 represent cDNA sequences encoding Beta vulgaris-like sugar
CC transport proteins

XX Sequence 1872 BP; 390 A; 579 C; 534 G; 369 T; 0 U; 0 Other;

Query Match 48.7%; Score 981.6; DB 7; Length 1872;

Best Local Similarity 77.4%; Pred. No. 3.4e-158;

Matches 1188; Conservative 1; Mismatches 345; Indels 0; Gaps 0;

QY 99 GCTTCGCGCGCTGCGGAGCGCTGCGCGCGGAGAGAGAGGCAAGCTCGCGTTCGCC 158
DB 128 GCTGCGCTCCGAGCGCGGGGCGAGTCCATCCAGGACAGGGCAATTTCAAGTACGCC 187
QY 159 TTGCGCTGCGCATCTCCCTCCATGACCTCCATCTCTCTCGGTACGATATCGGGTG 218
DB 188 TTCACCTGCGCGCTCTGTGCTTCCATGGCCACCATCGTCTCGGTACGACGTTGGGTG 247
QY 219 ATGAGCGGGCGCTGCGGTGTACATCAAGAGGACTTCAACATCAGTGAAGGAGTGGAG 278
DB 248 ATGAGCGGTGCGTGTGCTGTACATCAAGAGGAGCTGCAGATCACGAGCTGCAGTGGAG 307
QY 279 GTTCTCATGGGATACCTGAACTCTTACTGCTCATCGGCTCTTTCGCGCGGGGGAGCG 338
DB 308 ATCATGATGGCATCTGAGCGGTACGCGCTCATCGGCTCTTCTCGCGCGAGGAGCG 367
QY 339 TCGGACTGATCGCGCGGGGTACACCATCGTTCGCGCGCGCTCATATCTTTCGCGGG 398
DB 368 TCGGACTGCGCGCGCGCGGTACCGCTGCTTTCGCGCGCGCTCATATCTTTCGCGGG 427
QY 399 GSGTTCCTCATGGGTTTCGCGCTCAACTACGCGCATCTCATGTTTCGCGCGCTTCGTG 458
DB 428 TCCTTCTCATGGGTTTCGCGCTCAACTACGCGCATCTCATGTTTCGCGCGCTTCGTG 487
QY 459 GGCATCGGCTGCGGTACCGCTCATGATCGCGCGGTGATACCGCGAGGTTCGCGCG 518
DB 488 GGAATCGCGCTGCGGTACCGCATCATGTCGCGCGAGTGTACCGCGAGGTGTCGCGG 547
QY 519 GCGTTCGCGCGGTTCCTGATCGTTCGCGGAGGTTCATCAACTTCGCGCATCTGTCG 578
DB 548 GCGTTCGCGCGGTTCCTGATCGTTCGCGGAGGTTCATCAACTTCGCGCATCTGTCG 607
QY 579 CTCGGGTACGTCTCGAACTATGCTTTCCTCGCTTTCGCGCTGAACTTCGGGTGGCGCATC 638
DB 608 CTTGGGTACGTCTCGAACTATGCTTTCCTCGCTTTCGCGCTGAACTTCGGGTGGCGCATC 667
QY 639 ATGCTCGGATCGCGCGCGCGCTTCGCTGCTGCGCTCATGCTGCTTCGGCATTCGCG 698
DB 668 ATGCTCGGATCGCGCGCGCTTCGCTGCTGCGCTCATGCTGCTTCGGCATTCGCGCATC 727
QY 699 GAGTTCGCGCGGTGCTGTATGAAGGACCGCTTCGCGGACCGCAAGGTGTGTGTGAG 758
DB 728 GAGTTCCTCGTGGCTCGTATGAAGGCGCGCTTCGCGGACCGCAAGGCGCGCTTCGTG 787
QY 759 AGACCTCCGACACCGCGGAGGAGCGCGCGGAGCGCTTCGCGGACCGCAAGGCGCGCG 818
DB 788 AAGACCTCCGACACCGCGGAGGAGCGCGGAGCGCTTCGCGGACCGCAAGGCGCGCG 847
QY 819 GGCATCCCTGAGGAGCTCGACCGCGAGCTGTGACCGTCCCGCAAGAGGAGGAGCGGAAC 878
DB 848 GGCATCCCTGAGGAGCTTCGACCGCGAGCTGTGACCGTTCGCGGACCGCAAGGCGCG 907
QY 879 GAGAAGCGGTGTGGAAGAGCTCATCTTCGCGGACCGCGCGGAGCGCGCATTCGCGG 938
DB 908 GAGAAGCGGTGTGGAAGAGCTCATCTTCGCGGACCGCGCGGAGCGCGCATTCGCGG 967
QY 939 CTGTCGCGGATCGGCATCCACTTCTTCCAGCATGCGTTCGCGGAGTTCGCGTCTCTTC 998

Db 750 TTGGTGCAATACCTTTCGGTATCTCTTAACAGTAGAGAGTGTGGCGATGCCGGAGTCCCCAA 809
QY 709 GGTGCTGCTCATGAAGGAGCGCTTCGGCGAGCGCAAGTGTGTGGAGAGAACCTCCG 768
Db 810 GGTGCTTGTATGAGGGTCTGTTGGGAGAGCAAGAAAGTCTTAACAACAACTCAG 869
QY 769 ACAGCGGAGAGGCGCGGAGCGCTGGCGGACATCAAGGCGCGCGCGGATCTCCTG 828
Db 870 ACAGCAAGGAGAGCGCGGAGCGCTGGCGGAGCAAGCAAGCGCGGAGTCCCG 929
QY 829 AGGAGCTCGAGCGGAGCGCTGACCGTCCCAAGAGAGGAGCGGAAACGAGAAGCGGG 888
Db 930 AGAGTTGCAAGCAGCAGCTGCTCAGGTAATAACAAGCAAGCGTGAAG-----GTG 983
QY 889 TGTGGAAGAGCTATCTGTCGCGGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 948
Db 984 TATGGAAGAGCTTCTCTCTATCAAGCGCGGCAATTCGTCACATCGTAACTCGTCCG 1043
QY 949 TCGGATCCAGCTTCTCAGAGCGGTTGGGCACTTCACTCGTCTTCTACAGCGCTC 1008
Db 1044 TTGGTATTCATCTTCTCAAGCGTGGCGGTAGCGCGTCTTGTACAGCCCA 1103
QY 1009 TCGTGTTCAGAGCGCGGATTAAGCAAGCAACACATCTTGGGACCACTTGGCGGT 1068
Db 1104 GGATCTTCGAAAGCTGGATTAACAAGCAGCAGCATAAGCTTCTTCAACCGTGGCG 1163
QY 1069 TCGGTGTCACCAAGAGGCTTTCATCTGTGGGAGCTTCTTCAAGCGGCGTGGCG 1128
Db 1164 TTGGATTCGTTAAGACCGTGTCTTGTGGCGGTACGTTTACGTTGAGCGGTTGGTC 1223
QY 1129 GCGGCGCGTGTGGGAGCAGCGGCGGATAATCTCTCCCTCATCGGCTCGGCG 1188
Db 1224 GTGTCCTGTTGTTATGTTAGTGTGCGCGCATGGTCTCTCGTCTCAAGCTTCGA 1283
QY 1189 CCGGCTCAGCGTGTGGGAGCAGCGGCGGAGCAGGAGTACCTTGGGCGCATGGGCTAA 1248
Db 1284 TCAGCTCAGCTGTTATG---ATCATTCGAGAGGAAATTAATGTGGCGGTGGATCGA 1340
QY 1249 GCATCGCTCCAGCTCGCTACGTCGCTTCTTCTCCATCGGCTTGGCGGCTACGCT 1308
Db 1341 GCATAGCAGTGTGTTGCTTACGTCGCGGAGCTTCTTCACTCGGTCGCGGTCACGT 1400
QY 1309 GGGTATCAGTCGAGATCTCCGCTCCAGGTCGCGGCGTGGGTGTGCTGCTCGCG 1368
Db 1401 GGGTATAGTCTCAGATCTCCGCTTGGGCTGAGGCTGCGGCGCARGGTGCGGCGCGGAG 1460
QY 1369 TCGCGGCAACCGGTCACGAGCGGCTCATCTCCATGACCTTCTGCTGCTGTCGAGG 1428
Db 1461 TTGGGTTGAATAGGACCACTAGCGGCTTCTTCTTATTTGGGCAATGCTACTGTTGGTGGATAT 1520
QY 1429 CCATCACCATCGGCGGAGCTTCTTCTCTACTCGGCGCATCGCGGCTGCGCTGGGTGT 1488
Db 1521 CCATCAGTATTTGGTGGAGCTTCTTCTTATTTGGGCAATGCTACTGTTGGTGGATAT 1580
QY 1489 TCTTCTACACCTACTCCCGAGACCGCGCGGAGCGGAGCTCGAGGAGATG 1538
Db 1581 TCTTTTACACCGCTTTCCTGCTGAGACCGCGGAGAAACGCTCGAAGACATG 1630

RESULT 10

ABX93209
ID ABX93209 standard; cDNA; 1853 BP.

XX AC ABX93209;

XX DT 29-MAY-2003 (first entry)

XX DE cDNA encoding soybean sugar transport protein #3.

XX KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.

OS Glycine max.

XX PN US2002178468-A1.

XX PD 28-NOV-2002.

XX PF 17-JAN-2002; 2002US-00051902.

XX PR 24-APR-1998; 98US-0083044P.

XX PR 14-APR-1999; 99US-00291922.

XX PA (ALLE/) ALLEN S M.

XX PA (HITZ/) HITZ W D.

XX PA (KINN/) KINNEY A J.

XX PA (TING/) TINGEY S V.

XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

XX DR WPI: 2003-340957/32.

XX DR P-PSDB; ABU08337.

XX PT Novel plant sugar transport proteins and nucleic acid encoding the

XX PT protein useful for producing transgenic plants having altered levels of

XX PS sugar transport protein.

XX PS Claim 7; Page 34-35; 56pp; English.

XX CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABX93206-
CC ABX93211 represent cDNA sequences encoding Beta vulgaris-like sugar
CC transport proteins

XX SQ Sequence 1853 BP; 457 A; 434 C; 448 G; 513 T; 0 U; 1 Other;

Query Match 28.5%; Score 575.2; DB 7; Length 1853;

Best Local Similarity 63.6%; Pred. No. 6.7e-89;

Matches 909; Conservative 2; Mismatches 510; Indels 9; Gaps 2;

QY 109 CGCTCCGAGCGCTCGCGCGAAGAGAGGCAACGTCGCTTCGCTTCGCTCGC 168

Db 210 CACTTCAGATTTCGATCTCCAAAGAGCGCAAGAAAGCAAGTATGCTTTGTTGTG 269

QY 169 CCATCCTCGCTCCATGACCTCCATCTCTCGGTGATATATCGGTGATAGCGGG 228

Db 270 CTATGCTGCGCTCCATGACTTCCATCTTGTGTTATGATATTGAGTATGATGGAG 329

QY 229 CGTCTGTATCATCAAGAGGCTTCAACATCAGTACGAGGAGGAGGTTCTCATGG 288

Db 330 CAGCCATATACATAAAAGGGGACCTGAAAGTCTCGAGCAGCAAAATGAGATCTTGTCTG 389

QY 289 GCATCTGAACCTCTACTCGCTCATCTCGGCTCTTCCGCGCGGCGGCGGAGCTCGA 348

Db 390 GAATCATCAACCTATCTCTCTGATAGGCTCATGTCGCGGCGGAGAACCTCCGACTGA 449

QY 349 TCGGCGGCGGTACACCATCGTGTTCGCGCGCGGTATATCTTCCGCGGSGGTCTCTCA 408

Db 450 TAGTCTCCCGTTACAGATGTTTTCGCGCGCACCATCTTCTTGTCTCGAGCACTTCTCA 509

QY 409 TGGGGTTCCCGCTCAACTACGCAATGCTCATGTTCCGCGCGGTTCGTGCGCGCATCGCG 468

Db 510 TGGGTTTCTCCCGCAATATTCTTCTCATGTTTGGCGGTTCTGTCGTCGCAATGCA 569

QY 469 TGGGCTACGCTCATGATTCGCGCGCGGTGTACACCGCGGAGGTTCGCGCGGCTCGGCG 528

Db 570 TCGGCTACGCTCATGATAGCCCGGCTCTACACCGCGGAGTCTCCCGCGGCTCTCTC 629

QY 512 CTCGCGCGGTCGCGCGGTCGCTCTCTGACGTCGTTCCCGAGGTGTTTCATCACTTCGG 571
Db 558 TTCTCCTGCTCTCTCAAGAGGATTTCTACCTCTTTCTCTGAGGTTTTCATCAATGCAGG 617
QY 572 CATCTGCTCGGTAAGTCTCGAATATGCTTTCTCCGGCTTGCGGCTGAACCTCGGGTG 631
Db 618 AATTTTGTCTGATATATCTAATCACTTGTCAATTTTCAAGGCTCCCACTCATTTGAGCTG 677
QY 632 CGCATATGCTCGCATCGCGCGCGGCTCGCTGCTGCTCGCTGCTGCTGCTGCTGCTG 691
Db 678 GAGTTTATGCTCGAATCGGGCGATTCGAGCATATTTTGGCTATTGGAGTGTCTG 737
QY 692 CATCGCGAGTGC CGCGGCTGCTGTCATGAAGGAGCGCTCCGCGACGCCAAGTGCT 751
Db 738 TATGCTGAGTCAACCGCGGCTGCTGTTATGAGGAGGCTTGGAGATGCTAAGAGGT 797
QY 752 GCTGAGAGACCTCCGACGCGGAGGCGCGGAGCGCTGCGGACATCAAGGC 811
Db 798 GCTTAATCGCATCTGACTCGCTGAGGAGGCTCAACTTAGCTAGTGAGATTAAGCA 857
QY 812 CCGCGCGGCATCCTGAGGAGCTCGAGCGGAGCTGCTGACGTCGCTCCCAAGAGAGGAG 871
Db 858 GACAGCAGGAATCCAGCTGAGTGTGATGAAGACATCTATAAGGTTCAAAAAACAAGAT 917
QY 872 CGGAAACGAGAGCGGCTGGAAGAGCTCATCTGCTCCCGACCGCGGCATCGCGG 931
Db 918 AAAATCAGGGAATGAGTTTGGAAAGAGCTCTTCAACCCCTACTCCCGGAGTGAGCG 977
QY 932 CATCTGCTGCTCGGAGTCGCGATCCACTTCTCAGCATGCTGCTGGGCTATTCATCTCGT 991
Db 978 TCGAGTATTGAGTAICGGAATCACTTTTCCAGCAAGCTTCCGCGATCGAGCGCT 1037
QY 992 CGTCTTACAGCCCTCTGCTGTTTCAAGAGCGCGGATTAAGAACGACAAACACTTCT 1051
Db 1038 GGTTTTATACAGTCAAGGATCTTTCAGAGTGTGGAATCACAATGACGTAACAGT 1097
QY 1052 GGGCACCACTTGGCGGTCGCTGTCACCAAGAGCTTTTCACTGTTGTCGACTTTCTT 1111
Db 1098 ACTTCAACAGTAGCTGAGAGTGTTAAGACACTTTTCAATAGTTGCTACATTTCA 1157
QY 1112 CATGACGCGCTCGCGCGCGGCTGCTGCTGCGGAGCAGCGCGGAGATTCCTCTC 1171
Db 1158 ATTACAAAGTATGTAAGAGGCTTTACTATTAAAGTGTGCTGCTGCTGCTGCTGCTG 1217
QY 1172 CCTCATCGGCTCGCGCGGCTCACGCTGCTGCGGAGCAGCCGCGGAGATACC 1231
Db 1218 CATCTTAACCTTAGCCATGCTTTAACTGTTATTG---ATCCTCTCATCACAGATATC 1274
QY 1232 TTGGGCCATCGGCTTAAGCATCGCTCCACCTCGCCTAGCTGCTGCTTTCTTCCATCGG 1291
Db 1275 ATGGGCTATAGCTTGTATTAACCATGGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1334
QY 1292 CTTGGGCCCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1351
Db 1335 TTTAGGACCAATTACATGGGTATATAGTTACAGAGGTTTCCCTTAAAGTTAAGGCTCA 1394
QY 1352 GGGCTGCTCGCTCGCGCTCGCGGCAACCGGCTCACGAGCGGCTCATCTCCATGACCTT 1411
Db 1395 AGGTACTAGCATGGTGTGGCTGTTATAGAGTTGTAAGTGTGTTATTTCAATATTTT 1454
QY 1412 CTTGCTGCTGCTCAAGGCTATCAGCATCGGCGGAGCTTTCTCTCTACTCCGCGATCGC 1471
Db 1455 CTTGCTGCTGCTCAATAGATTACAGAGTGTGCTGCTTCTTTTGTTCGAGGAGTTCG 1514
QY 1472 CGGCTGCTGCTGCTGCTTCTTCTACACTCTCCGAGACCGCGCGGAGCTGGA 1531
Db 1515 TATTAATGCTGCTGCTTCTTCTTCTGAGCTTTCTTCTGAGACTAGAGGCGGTACACTTGA 1574
QY 1532 GGAGATGAGCAAGCTGTTTCGCGGA 1555
Db 1575 GAAATATGATGATGATGTTTGAAGA 1598

RESULT 12
AAC43261
ID AAC43261 standard; DNA; 1527 BP.
XX
AC AAC43261;
XX
DT 17-OCT-2000 (first entry)
XX
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38621.
DE
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01213180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0128264P.
PR 29-MAR-1999; 99US-0128785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130445P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135621P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.


```
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139730P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140333P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140685P.
PR 28-JUN-1999; 99US-0140833P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142300P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144338P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.

PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 05-OCT-1999; 99US-0157117P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 23.3%; Score 470.2; DB 3; Length 1527;
Best Local Similarity 59.4%; Pred. No. 5.2e-71;
Matches 834; Conservative 1; Mismatches 559; Indels 9; Gaps 2;

QY 152 GTTCGCTTCGCTCGCCATCCTCGCTCCATGACCTCCATCCTCGGCTAGCATAT 211
Db 60 GTTCGCTTCGCTCGCCATCCTCGCTCCATGACCTCCATCCTCGGCTAGCATAT 119
QY 212 CGGGGTGATGACGGGGGTGCTGCTATACATCAAGAGGACTTCAACATCAGTACGGGAA 271
Db 120 GGGAGTTATGACGGGAGCTCAGATATTTATAAGAGATGATCTAAATAAAGCACTCA 179
QY 272 GGTGAGGTTCTCATGGGCATCTGAACCTCTACTCGCTCATCGGCTCTTCGCGGGGG 331
Db 180 GATTGAGTTTGGCCGGNACTTGAATCTTTGTGCACTCGTCGATCACTAACGGGGG 239
```


Db 264 AATAGGAGTTCTAGTCGAATCTTAAACATCTACTGCTCTTCGGCTCTTTCGAGCTGG 323
QY 332 GCGGAGCTGCGAGTGAATCGCGCGCGGTGACCAATCGTGTTCGCGCGGTGATATCTTT 391
Db 324 TAGAACATCCGATTTGGATCGGACGAGTTACATATCGTGTAGCTGGTGAATCTCTTT 383
QY 392 CGCGGGGGTTCCTCATGGGTTTCGCGGTCAACTAGCCATGCTCATGTTCGCGCGCTT 451
Db 384 GTTGGCGCACTACTCAATGGTTTCGCCAACAATACGCAATTCCTTAGGTAGGCCGTTT 443
QY 452 CGTGCGCGCATCGCGGTGGCTAGCGCTCATGATCGCGCGGTGTACACCGCGAGGT 511
Db 444 TGTAAACAGGAATAGTGTGTGTTATGCACTTATGATCGACCTGTTTACATCGAGAGT 503
QY 512 GTCCCGCGCTCGCGGTGCTTCCTGACGCTGTTCCGAGGTGTTCATCACTTCGG 571
Db 504 TTCTCTCTCTTCAAGAGGATTTCTTACCTCTTTTCCCTGAGGTTTTCATCAATGAGG 563
QY 572 CATCTGCTCGGTACGCTCTCGAACTATGCTTTCTCCGCTTGCGGCTGAACTTCGGGTG 631
Db 564 AATTTGCTTGAATATATCTATCTCGCATTTTCAAGCTTCCAACTCATTTGAGCTG 623
QY 632 CGCATCATGCTCGCATCGCGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 691
Db 624 GAGTTTATGCTCGGAATCGGGGATTTCCGAGCATATTTTGGCTATTGAGGTGCTTGC 683
QY 692 CATCGCGAGTCCGCGGTGCTGCTCATGAAGGAGCTTCGCGGACGCCAAGGTGT 751
Db 684 TATGCTGAGTACCGAGTGGCTGTTATGACGGAAGCTTGGAGATGCTAAGAAGT 743
QY 752 GCTGGAAGACCTCCGACACGCGGAGGCGCGCGAGCGCTGCGCGCATCAAGG 811
Db 744 GCTTAATCGAATATCTGACTCGCTGAGAGGCTCAACTTAGGCTCAGTGAGATTAAGCA 803
QY 812 CGCGCGCGCATCTGAGAGCTCGAGCGAGCTGCTGACCTGCCAGAGAGGAG 871
Db 804 GACAGAGGAATCCAGCTGAGTGTGATGAAGACATATAGAAGTTGAAAAAACAAGAT 863
QY 872 CGAAACGAGAGCGGTGTGAAGAGGCTCATCTGCTCCCGACCCCGGCTGCGGCG 931
Db 864 AAATCAGGGAATCGAGTTGGAAGAGCTTCTTCAACCTACTCCCGCAGTGAGCG 923
QY 932 CATCTGCTGCTCGGATCGCATCCACTTCTCCAGCATGCGTGGGCAATCACTCGT 991
Db 924 TGCAATGATTCAGATATCGGAATTCATTTTCCAGCAAGCTTCCGCGATCGACGCGT 983
QY 992 CGTCTCTACAGCCTCTCGTGTCAAGAGCCCGGATTAAGGAACGACAAACACTTCTT 1051
Db 984 GTTTTATACAGTCAAGGATCTTTCAGTGTGCTGGAATCACAAATCGATTAACAGTT 1043
QY 1052 GGGCACCACTTGGCGGTTCGCTGTCACCAAGAGCTTTTTCATCTGTTCGCACTTCTT 1111
Db 1044 ACTTGCAACAGTAGCTGTAGAGTTGTTAAGACACTTTTCATATTAGTTGCTACATTCA 1103
QY 1112 CATCGAGGCTCGGCGCGCGCTGCTGCTGGGACGACGCGCGGATATCTCTC 1171
Db 1104 ATTAGCAAGTATGTAAGAGGCTTTTACTATTAAAGTGTGTGTTGATGATCATAGC 1163
QY 1172 CCTCATCGGCTCGGCGCGGCTCACGCTGCTGCGGACGACCCCGCAGATAC 1231
Db 1164 CATCTTAACCTTAGCCATGCTTTTAACTGTTATG---ATCACTCTCATACAAGATTAC 1220
QY 1232 TTGGCCATCGGCTAAGCATCGCTCGACCTCGCTAGTCGCTTCTTCTCCATCGG 1291
Db 1221 ATGGGCTATACCTTGTGTATAACCATGGTTGTGCTGTGCTGCTGCTGCTGCTGCTG 1280
QY 1292 CTTTGGCCCATCAGCTGGGTGTACGCTGGAGATCTTCCCGCTCCAGGTGCGCGGCT 1351
Db 1281 TTTAGGCAATATACATGGGTATATAGTTAGAGGTTTTCCTTGGTTAGGCTCA 1340
QY 1352 GGGTCTGCTGCGGTGCGCGCAACCGGCTCACAGCGGCTCATCTCCATGACCTT 1411

Db 1341 AGGTACTAGCATGGGTGCTGCTTTAATAGAGTTGTAAGTGGTATTATTTCAATATTTT 1400
QY 1412 CTGTCTGCTTCCAGGCCATCACCATCGCGGAGCTTCTTCTCTACTTCGGCAATCGC 1471
Db 1401 CTTGCTCTTTTGCATACATAAGATTACACAGGTGCTGCTTCTTTTCTCGGAGGATTC 1460
QY 1472 CGGCTCGCTGCTGCTGCTTCTTCTACACCTACTTCGCGGAGACCGCGGCGGACGCTGGA 1531
Db 1461 TATTATGCTGCTGCTTCTTCTTCTGAGCTTCTTCTGAGCTAGAGGGGCTACACTTGA 1520
QY 1532 GGAGATGAGCAAGCTTTCGCGGA 1555
Db 1521 GAATATGCTAGTGTGTTTGAAGA 1544

RESULT 14

ABL41878/c
ID ABL41878 standard; DNA; 1662 BP.

XX ABL41878;

XX AC ABL41878;

XX DT 11-JUN-2002 (first entry)

XX Nucleotide sequence of a linear polyol transporter.

XX Linear polyol; mannitol; polyol transporter; carbon source; plant;

XX pathogen resistance; salt stress; sorbitol; dulcitol; galactitol;

XX inositol; ribitol; xylitol; ss.

XX Arabidopsis thaliana.

XX WO200204647-A1.

XX 17-JAN-2002.

XX 22-JUN-2001; 2001WO-FR001979.

XX 11-JUL-2000; 2000FR-00009032.

XX (CNRS) CENT NAT RECH SCI.

XX Lemoine RRP, Noiraud NEJ;

XX WPI; 2002-154933/20.

XX New polyol transporter protein from plants, for selecting transformed cells and for imparting pathogen and salt-stress resistance to plants.

XX Disclosure; Page 44-45; 66pp; French.

CC The present sequence encodes a polypeptide which is a transporter of a linear polyol. DNA encoding linear polyols is used to produce a selection system for transformed cells, based on the use of polyols as the only carbon source. It is also used to produce transgenic plants with

CC increased resistance to pathogens and salt stress. The use of a linear polyol for selection eliminates the need for toxic selection reagents such as antibiotics. Also, the linear polyol is not essential for the plant, once selection has been made. The linear polyol has a main chain of 5-8, preferably 6, carbon atoms and is selected from sorbitol,

CC dulcitol, galactitol, (myo-)inositol, ribitol, or xylitol.

XX Sequence 1662 BP; 508 A; 445 C; 306 G; 403 T; 0 U; 0 Other;

Query Match 22.0%; Score 443.2; DB 6; Length 1662;

Best Local Similarity 59.0%; Pred. No. 2.1e-66;

Matches 798; Conservative 1; Mismatches 544; Indels 9; Gaps 2;

QY 203 CTACGATATCGGGGTATGAGCGGGGCTGCTGTACATCAAGAGGACTTCAACATCAG 262

Db 1417 CTTAGATCGGAGTTATGAGCGGAGCTCAGATATTATAGAGATGATCTMAAATAAA 1358

QY 263 TCACGGGAGGTGGAGGTTCTCATCGGCATCTAACCTCTACTCGCTCATCGGCTCCTT 322

Db 1357 CGACACTCAGATTGAAGTTTGGCGCGGAATCTTGAAATTTTGTGCACTCGTCGGATCACT 1298
QY 323 CGCGCGGGCGGACGTCGAGCTGATCGCGCGCGGTACACCATCGTGTGTTGCGCGCGGT 382
Db 1297 AACGGCGGGAAAGCTCTGACGTCATCGCGCGAGCTTACACCATAGCTCTCTCGCGGT 1238
QY 383 CATATCTTCGCGGGGSGTCTCATGCGGTTGCGCGTCAACTAGCGCATGCTCATGTT 442
Db 1237 GATATCTTAGTGGGCTCAGTCTTATGCGTTACGCGCCAACTATCTGTTTGTGATGGT 1178
QY 443 CGCGCGCTTCGTCGCGCATCGCGTGTAGCGCTCATGATCGCGCGGTGTACAC 502
Db 1177 CGAGCATGTATGCGCGAGTTGTGTAGGTTGTCTCATGATGCTCGGTTTACTC 1118
QY 503 CGCGAGGTGTGCGCGGCTCGCGCGGTGCTTCTGACGTGCTTCGCGGAGGTGTTCAAT 562
Db 1117 TGCGGAGATATCTCTGCGTCACATAGAGGCTTTCTCACTTCTCTACCTGAGCTTTGAT 1058
QY 563 CAACTTCGGCATCTGCTCGGTGAGTCTCGAATCATGCTTTCTCCGCTTCCGCTGAA 622
Db 1057 TAGTCTTGCGATTTTACTAGGCTATGCTCGAATTAATCTGTTTGGAACTGACGTTGAA 998
QY 623 COTCGGGTGGCGCATCATGCTCGGCATCGCGCGCGCGCTCCGTCGTCTCGCGCTCAT 682
Db 997 ACTTGATGGAGATGATGCTTGGAAATCGACGCTTCCGCTCTTATATTGGCTTTGG 938
QY 683 GGTGTCGGCATCGCGGAGTCCCGGTGCTGCTCATGAGGAGCGCTTCGCGGAGC 742
Db 937 GATTACAAGATGCGCGAGTCCCGAGTGTGATGCAAGTAGGTAGTAGAAGC 878
QY 743 CAAGTGTGTGGAAGAAGCTCCGACACCGCGGAGGCGCGGAGCGCTCGCGCA 802
Db 877 CAAGAAATATGTTTGGTATCAATACGGAAGAAGCTGAAGAAGCTTCAGAGA 818
QY 803 CATCAAGCGCGCGCGCATCCCTGAGAGCTCGACGCGACGTGCTGACGCTCCCA 862
Db 817 CATCTTAACCGCTCGGGAAGTAGATGTACAGAGATAAAGAAAGTTGTGCGCGGTAA 758
QY 863 GAGAGGAGCGAAACGAGAGCGGTGTGAAGAGCTCATCTGCTCCGACGCGCGGC 922
Db 757 GA---AGAAATACGGTAAAGTTTGGAGGAGTTGTGATAAACCAGCACTGC 701
QY 923 CATGCGCGCATCTCTGCTTCGCGGATCGGCATCCATCCATCTTCAGCATGCTTGGGAT 982
Db 700 GGTGCTTTGATCTTGTATGACGCGCTTGGGATACATCTTTTCGAGCATGCGAGCGGAT 641
QY 983 TCACTCGGCTCTTCTACAGCGCTCTGCTGTCAAGAGCCCGGATTAACGACGAA 1042
Db 640 TGAGGCGGTCGTTGTACAGCGCGAGAAATTTCAAGAAAGCGCGAGTTGTTTCAAGA 581
QY 1043 ACATCTTTTGGCACCACCTTGGCGCTTGGTGTACCAAGAGGCTTTTCATCTTTGGC 1102
Db 580 CAAGCTCTTACTAGCACGCTCGGTAGTTTAACTAAGGCTTTTTCATAATAATAGC 521
QY 1103 GACTTTCTTCATCGAGCGGCTCGCGCGGCGGCTGTGCTGGCACACGCGCGGAT 1162
Db 520 CACTTTCTTGTCTGACAGGTTAGTCTGTAGAAAGCTCTGTGTACCAATGCGGTTGGAT 461
QY 1163 AATCTCTCTCTCATCGGCTCGCGCGGCTCACCGTCTCGCGGAGCAGCAGCAGC 1222
Db 460 GGTCTTCGATTTAGCGTTTACGCTTGTAGTCTACGATGTT-----TCAGCATTTGG 407
QY 1223 CAAGATACCTTGGGCAATCGGCTTAAGCATCGCTTCACCTCGCTCGCTAGCTGCTTCTT 1282
Db 406 GCGGTTAGCGTGGCAATTGAGTTTAAAGCATCGCTCTACGATGCAATTTGTAGCTTTT 347
QY 1283 CTTCATCGGCTTGGCCCATCAGCTGGGTGTACAGCTCGGAGATCTTCCGCTCCAGT 1342
Db 346 CTCCATAGGCTTGGACCTATACATGGGTGTACAGCTCCGAGATTTCCGCTTGAGCT 287
QY 1343 GCGCGGCTGGGCTGCTCGCTCGGCGTCCGCCAACCGGCTACACAGCGGCTCATCTC 1402
Db 286 TAGGGCTCAGGAGCGAGTATAGTGTGTTGCGGTTAATAGGATCATGAACGCGAGCGTCTC 227

QY 1403 CATGACCTTCTGTGCTGTCCAGGCCATCAACATCGCGGAGCTTCTTCTCTACTC 1462
Db 226 AATGAGTTTTTCTGTGATGACGAAGCGGATACGACAGGGGCGGTGTTCTTCTATTG 167
QY 1463 CGSCATCGCGCGCTCGCTCGGTGTTCTTCTACACTACCTCCGCGAGACCGCGCG 1522
Db 166 CGGAATAGCGGTGCGGCATGTTGTTCTTCTTATGTTGCGGAGACGAAGGAT 107
QY 1523 GAGCTGGAGGAGATGAGCAAGCTGTTGCGG 1554
Db 106 GCCATTGGAGGATGGAGAGCTTTTGGTG 75

RESULT 15

ABZ12990
ID ABZ12990 standard; DNA; 1482 BP.

XX AC ABZ12990;

XX DT 21-JAN-2003 (first entry)

XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 795.

XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX OS Arabidopsis thaliana.

XX PN WO200216655-A2.

XX PD 28-FEB-2002.

XX PF 24-AUG-2001; 2001WO-US026685.

XX PR 24-AUG-2000; 2000US-0227866P.

XX PR 26-JAN-2001; 2001US-0264647P.

XX PR 22-JUN-2001; 2001US-030011P.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PA Harper JF, Kreps J, Wang X, Zhu T;

XX PI WPI; 2002-304127/34.

XX PT Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

XX PS Claim 144; SEQ ID NO 795; 577bp + Sequence Listing; English.

XX CC The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

XX SQ Sequence 1482 BP; 353 A; 311 C; 391 G; 427 T; 0 U; 0 Other;

Query Match 21.5%; Score 433.8; DB 6; Length 1482;
Best Local Similarity 57.8%; Pred. No. 8.4e-65;
Matches 809; Conservative 1; Mismatches 583; Indels 6; Gaps 2;

QY 151 GGTTCGCTTCGCTCGGCGCATCTCGCTCCATGACCTCCATCTCTCCGCTACGATA 210
Db 44 GATTGCTCTCAATGTGCTATCGCTCCATCTCTCCATCATCTTTGGTTACGATA 103

Qy 211 TCGGGGTGATGAGCGGCGCTCGCTGTACATCAAGAGGACCTCAACATCAGTCAAGGGA 270
Db 104 CTGGTGTATGAGTGGAGCGATGCTGTTTATAGAGAGATTTGAAGACAACGACGTTT 163
Qy 271 AGGTGGAGTTCATCGGCGATGACTGAACTCTACTCGCTCATCGGCTCCTTCGGGCGG 330
Db 164 AAATAGAGTTCCTCACTGGAATTTCTCAACCTTTGTGCCCTTGTGCGATCATTCCTCGCG 223
Qy 331 GCGGAGCTCGGACTGATCGGCGGCGGTACACCATCGTGTGTTGCGCGCGCTCATATTTCT 390
Db 224 GAAGAACTCGGACATTAATCGGACGGGTACACATCGTCTTGGGCTCAATACTATTCA 283
Qy 391 TCGCGGGGGGTTCCTCATGGGTTCCCGGTCAACTAGCCATGCTCAATGTTCCGCGCT 450
Db 284 TGTTAGGCTCAATATTGATGGTGGGTCGGAATTTATCCGTTTCTCCTATCCGGTAGAT 343
Qy 451 TCGTGGCGGCGATCGGCGGTGGGCTAOCGCTCATGATCGCGCGGTGTACACCGCGGAG 510
Db 344 GCACCGCTGGAATCGGAGTCGTTTCTGCTGTGATGGTGTCTCGGTTTACTCTGCCGAGA 403
Qy 511 TGTCCCGCGGTCCGCGGTGGGTTCTGACGTGTTTCCCGAGGTGTTTCATCAACTTCG 570
Db 404 TCGCAACTGCTTCACATAGAGGACTTTAGCTTCTCTTCTCCTCACTTTGTATCAGTATAG 463
Qy 571 GCATCCTGCTCGGTACGCTCGAATCATGCTTCTCCCGCTTCCGCTGAACCTCGGT 630
Db 464 GGATTTTACTAGGTTATCTGTAATTAATCTTCTCCAACTTACCTATGATCATCGGTT 523
Qy 631 GCGCATCATGCTCGGCGATCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 690
Db 524 GGAGACTCATGCTCGGTATAGCGCGGTTCCGTCCTTAGTGTAGCGTTCGGGATCTTGA 583
Qy 691 GCATGCGGAGTCCCGCGGTGCTGTGATGAAAGGACGCTCGCGGAGCCCAAGGTGG 750
Db 584 AAATGCGGGAATCTCCACGGTGTGATTAAGAAAGCGGCTTAAAGAAAGGCAAGGAGA 643
Qy 751 TGCTGGAAGAGACTCCGACACGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 810
Db 644 TATTGGAATTTGATCTAATTTCCCTGAGAGAGAGAACTCCGCTTCAAGACATCAAG 703
Qy 811 CCGCGCGCGGATCCCTGAGGAGCTCGACGCGGAGCTGTTGACCGGTCGCCCAAGAGAGGA 870
Db 704 CTGCTCGGGAATCGACCCGAAATCGGTAGACGATGTTGTGAAATGG---AGGGTAAGA 760
Qy 871 CGGAAACGAGAGCGGCTGAGAGAGCTCATCTGTCCCGACCCCGGCGGCGGCGGCGGCG 930
Db 761 AGACTCATGTGAAGAGTGTGAAGAGCTCATCTTAAGACCAACTCTGACAGTGAAGAC 820
Qy 931 GCATCCTGCTGCCGGATCGGATCCACTTCTTCCAGCATGCGTGGGCAATTCAGTCCG 990
Db 821 GTGTTCTTTAACTGCTCTTGGGATTCATTTCTTCAACAGCTCCGGATCGAGCAG 880
Qy 991 TCGTCTCTACAGCCCTCTCGTGTTCAGAGCCCGGATTAAGAGCAACAACACTTCT 1050
Db 881 TGCTTTTATACGTTCCGAGGATCTTTAAGAAAGCAGGAATCAGCACTAAAGACAAGCTTT 940
Qy 1051 TGGGCAACACTTGGCGGTGCTGCTGCTACCAAGAGGCTTTTCACTGTTGCGGACTTCT 1110
Db 941 TCTTGGTTACTATCGGTGTGGAATCATGAAACGACGTTTATTTTCACTGCGACTTAT 1000
Qy 1111 TCATCGAGCGGCTCGGCGCGGCGGCTGTTGTGTCGAGCAGCGGCGGATTAATCCTCT 1170
Db 1001 TGTTAGACAAGGTAGTTCGAAGGAAGCTTTTGTGACCAAGCTTGGAGGAATGTCATTC 1060
Qy 1171 CCCTCATCGGCTCGGCGCGGCTCACCGTCTGTCGCGGAGCACCCCGGCGGAGATAC 1230
Db 1061 CGTTGACAATGTTGGGATTTGGGCTTAC---AATGGCCCCAAAATGCTGGCGGGAATTAG 1117
Qy 1231 CTTGGGCGCATCGGCTAAGCATCGCCTCCACCTCGGCTACGTCGCTTCTTCTCCATCG 1290
Db 1118 CGTGGGCTTTAGTACTGAGCATAGTTGGGCTTATAGTTTCGTGGCGTTTTTCTCTATTG 1177
Qy 1291 GCCTTGGCCCCATCACTGGTGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGGCGC 1350

Db 1178 GGCTCGGCCCAATAACTTTGGGTCTACAGTTCTAGGTTTCCCGTTGAAGCTTAGGGCAC 1237
Qy 1351 TGGGCTGCTCGCTCGGCTCGCGGCAACCGCGTCACCAGGGCGTCACTCCATGACCT 1410
Db 1238 AAGGAGCGAGTCTCGGCGTTCGGGTGAACAGAGTAATGAACGCCCGTGTGATGCT 1297
Qy 1411 TCTGTGCTGTGTCGAAGGCCATCACATCGGCGGCGAGCTTCTTCTTACTCCGCGATCG 1470
Db 1298 TTTTGTGCTTGACTAGTGGGATAACACCGGTGGAGCTTCTTTATGTTCCGCGAGTTG 1357
Qy 1471 CCGCGCTCGCCTGGGTGTTCTTCTACACTACCTCCGAGACCCGCGCGGAGCGCTGG 1530
Db 1358 CGGCACTGGCGTGGAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1417
Qy 1531 AGGAGATGAGCAAGCTGTT 1549
Db 1418 AAGAAATCGAAGCGCTTT 1436

Search completed: July 2, 2004, 02:47:01
Job time : 1163 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 02:27:34 ; Search time 173 Seconds
(without alignments)

Title: US-10-051-902A-21

Perfect score: 100
Sequence: 1 cttacatgtaagttcgtgcc.....aaaaaaaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```
Database : Issued Patents_NA.*
1: /csgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /csgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /csgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /csgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /csgn2_6/ptodata/2/ina/PTUTS_COMB.seq.*
6: /csgn2_6/ptodata/2/ina/backfiles1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2016.6	100.0	2017	4	US-09-291-922-21	Sequence 21, Appl
2	1094.6	54.3	2089	4	US-09-291-922-25	Sequence 25, Appl
3	1005	49.8	1914	4	US-09-291-922-19	Sequence 19, Appl
4	981.6	48.7	1872	4	US-09-291-922-27	Sequence 27, Appl
5	575.2	28.5	1853	4	US-09-291-922-23	Sequence 23, Appl
6	343.6	17.0	615	4	US-09-291-922-17	Sequence 17, Appl
7	188.8	9.4	1675	4	US-09-678-686B-1	Sequence 1, Appl
8	168.4	8.3	1776	4	US-09-678-686B-11	Sequence 11, Appl
9	147.2	7.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl
10	145.4	7.2	1545	4	US-09-488-039A-4731	Sequence 4731, Ap
11	141.2	7.0	1752	4	US-09-678-686B-17	Sequence 17, Appl
12	136	6.7	1167	4	US-09-678-686B-9	Sequence 9, Appl
13	138	6.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
14	118	5.9	1431	4	US-09-488-039A-4762	Sequence 4762, Ap
15	102	5.1	1344	4	US-09-252-991A-3992	Sequence 3992, Ap
16	102	5.1	1479	4	US-09-252-991A-3960	Sequence 3960, Ap
17	96.6	4.8	1589	4	US-10-162-012-45	Sequence 45, Appl
18	96.6	4.8	4385	4	US-10-162-012-43	Sequence 43, Appl
19	91.2	4.5	1320	4	US-09-252-991A-10535	Sequence 10535, A
20	91.2	4.5	1473	4	US-09-252-991A-10639	Sequence 10639, A
21	89.4	4.4	1407	4	US-09-252-991A-11395	Sequence 11395, A
22	89.4	4.4	2871	4	US-09-252-991A-11431	Sequence 11431, A
23	88.4	4.4	2601	4	US-09-291-922-7	Sequence 7, Appli
24	87.4	4.3	2824	4	US-09-291-922-1	Sequence 1, Appl
25	87.4	4.3	2856	4	US-09-643-597-135	Sequence 135, App
26	87.4	4.3	2856	4	US-09-488-884A-135	Sequence 135, App
27	87.4	4.3	2856	4	US-09-543-615A-135	Sequence 135, App

ALIGNMENTS

RESULT 1

US-09-291-922-21
; Sequence 21, Application US/09291922

; Patent No. 6383776

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plan

FILE REFERENCE: BB-1163

; CURRENT APPLICATION NUMBER: US/

; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 50,083,044

EARLIER APPLICATION NUMBER: 60/083,
EARLIER FILING DATE: April 24, 1998

; EARLIER FILING DATE: AP.
: NUMBER OF SEC ID NOS. 3

NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97

; SOFTWARE: MLC
: SEO ID NO 21

; LENGTH: 20

```

; TYPE: DNA

```

; ORGANISM: Oryza sativa

US-09-291-922-21

Query Match	Score	2016.6
100.0%	Score	2016.6

Best Local Similarity 100.0%; Pred. No. 0;

QY	1	CTTACATGAAGCTCGTGC	CGGCACGAGCTTACACT	CGACCGCCACTCTGTAC	ACGGCC	60
Db	1	CTTACATGAAGCTCGTGC	CGGCACGAGCTTACACT	CGACCGCCACTCTGTAC	ACGGCC	60
QY	61	CAGACGAGCCTCTCTCT	CTCTGCACACCGAGATG	GCTTCCGCCGCGCTGCC	GGAG	120
Db	61	CAGACGAGCCTCTCTCT	CTCTGCACACCGAGATG	GCTTCCGCCGCGCTGCC	GGAG	120
QY	121	CGCTCGCGCGAAGAAAG	GGCAAGTCGCGCTTCG	CGCTCGCGCATCTCTCG	CGCT	180
Db	121	CGCTCGCGCGAAGAAAG	GGCAAGTCGCGCTTCG	CGCTCGCGCATCTCTCG	CGCT	180
QY	181	CCATGACCTCCATCTCT	CTCGGCTACGATTCGG	GGGTGATGACGGGGGG	TGCTGTAC	240
Db	181	CCATGACCTCCATCTCT	CTCGGCTACGATTCGG	GGGTGATGACGGGGGG	TGCTGTAC	240
QY	241	TCAAGAGGATTCAACAT	CAGTCACGGGAAGTGG	AGGTTCTCATGGGCAT	ACTGAAC	300
Db	241	TCAAGAGGATTCAACAT	CAGTCACGGGAAGTGG	AGGTTCTCATGGGCAT	ACTGAAC	300
QY	301	TCTACTCGCTCATCGG	CTCCTCCGCGCGGGG	CGGACGTCGACTGG	ATCGGCCGCGGT	360
Db	301	TCTACTCGCTCATCGG	CTCCTCCGCGCGGGG	CGGACGTCGACTGG	ATCGGCCGCGGT	360

QY 361 ACACCATCGTGTTCGCCCGCGTCATATTTTCGGGGGGSGTTCCTCATATGGGTTCGCCG 420
DB 361 ACACCATCGTGTTCGCCCGCGTCATATTTTCGGGGGGSGTTCCTCATATGGGTTCGCCG 420
QY 421 TCAACTAGCCCATGCTCATGTTCGCGCGCTTCGTGGCGCGCATCGCGTGGGCTACCGCG 480
DB 421 TCAACTAGCCCATGCTCATGTTCGCGCGCTTCGTGGCGCGCATCGCGTGGGCTACCGCG 480
QY 481 TCATGATCGCGCGGTGTACACCGCGAGGTGTCCCGCGCGTCCGCGGTTCCTGA 540
DB 481 TCATGATCGCGCGGTGTACACCGCGAGGTGTCCCGCGCGTCCGCGGTTCCTGA 540
QY 541 CGTCGTTCCCGAGGTGTTCATCAACTTCGGGCATCCTGCTCGGGTACGTTCTCGAACTATG 600
DB 541 CGTCGTTCCCGAGGTGTTCATCAACTTCGGGCATCCTGCTCGGGTACGTTCTCGAACTATG 600
QY 601 CTTTCTCCCGCTTCGCGCTGAACTTCGGGTGGCGATCATGCTCGGCATTCGGCGCGCGC 660
DB 601 CTTTCTCCCGCTTCGCGCTGAACTTCGGGTGGCGATCATGCTCGGCATTCGGCGCGCGC 660
QY 661 CGTCGCTGCTGCTCGCGCTCATGTGCTCGGCATCGCGAGTTCGGCGGTGGTGTCA 720
DB 661 CGTCGCTGCTGCTCGCGCTCATGTGCTCGGCATCGCGAGTTCGGCGGTGGTGTCA 720
QY 721 TGAAGGAGCGCTTCGCGAGCGCAAGGTGTGCTGGAGAGACCTTCGACACGCGCGAGG 780
DB 721 TGAAGGAGCGCTTCGCGAGCGCAAGGTGTGCTGGAGAGACCTTCGACACGCGCGAGG 780
QY 781 AGGCGCGGAGCGCTTCGCGAGCATCAAGGCGCGCGCGCATCCCTGAGGAGTTCGACG 840
DB 781 AGGCGCGGAGCGCTTCGCGAGCATCAAGGCGCGCGCGCATCCCTGAGGAGTTCGACG 840
QY 841 GCGAGCTGGTACCGTCCCAAGAGAGGAGCGGAACGAGAAGCGGTGTGGAAGGAGC 900
DB 841 GCGAGCTGGTACCGTCCCAAGAGAGGAGCGGAACGAGAAGCGGTGTGGAAGGAGC 900
QY 901 TCATCTGTCCCGAGCGCGCGCATTCGCGCATCTGCTCGGGATCGGCATCCACT 960
DB 901 TCATCTGTCCCGAGCGCGCGCATTCGCGCATCTGCTCGGGATCGGCATCCACT 960
QY 961 TCTTCAGCATCGTGTGGGCACTTCACTCCGCTCTCTACAGCCCTCTCGTGTCAAGA 1020
DB 961 TCTTCAGCATCGTGTGGGCACTTCACTCCGCTCTCTACAGCCCTCTCGTGTCAAGA 1020
QY 1021 GCCCGGATTAACGAACGACAAACACTTCTTGGGACCACTTGGCGGTTGGGTGCACCA 1080
DB 1021 GCCCGGATTAACGAACGACAAACACTTCTTGGGACCACTTGGCGGTTGGGTGCACCA 1080
QY 1081 AGAGGCTTTTCACTCTTGTGGGCACTTCTTTCATCGAGCGGCTCGGGCGCGCGGTGT 1140
DB 1081 AGAGGCTTTTCACTCTTGTGGGCACTTCTTTCATCGAGCGGCTCGGGCGCGCGGTGT 1140
QY 1141 TGCTGGGAGCAGCGCGGAGTAATCTCTCCCTCATCGGCTTCGGCGCGCGGTTCACCG 1200
DB 1141 TGCTGGGAGCAGCGCGGAGTAATCTCTCCCTCATCGGCTTCGGCGCGCGGTTCACCG 1200
QY 1201 TCGTCGGCGAGCAGCGCGAGATACCTTGGCCATCGGCTTAAGCATCGCCTCCA 1260
DB 1201 TCGTCGGCGAGCAGCGCGAGATACCTTGGCCATCGGCTTAAGCATCGCCTCCA 1260
QY 1261 CCTCGGCTAGCTCGGCTTCTTCTCCATCGGCTTGGGCCCATCACGTGGGTGTACAGCT 1320
DB 1261 CCTCGGCTAGCTCGGCTTCTTCTCCATCGGCTTGGGCCCATCACGTGGGTGTACAGCT 1320
QY 1321 CGGAGATCTTCCGCTCAGGTTCGGCGCTGGGCTGCTGCTCGGCTCGCCCGCAACC 1380
DB 1321 CGGAGATCTTCCGCTCAGGTTCGGCGCTGGGCTGCTGCTCGGCTCGCCCGCAACC 1380
QY 1381 GCGTCACGAGCGGCTCATCTCCATGACCTTCTGCTGTCAAGGCCATCACCATCG 1440
DB 1381 GCGTCACGAGCGGCTCATCTCCATGACCTTCTGCTGTCAAGGCCATCACCATCG 1440

QY 1441 GCGCAGCTTCTTCTTCTACTCCGGCATCGCGCTCGCCTGGGTGTCTTCTACACT 1500
DB 1441 GCGCAGCTTCTTCTTCTACTCCGGCATCGCGCTCGCCTGGGTGTCTTCTACACT 1500
QY 1501 ACCTCCCGAGAGCCCGCGCGGAGCGCTGGAGAGATGAGCAAGCTTTCGGGACACGG 1560
DB 1501 ACCTCCCGAGAGCCCGCGCGGAGCGCTGGAGAGATGAGCAAGCTTTCGGGACACGG 1560
QY 1561 CGCGCCCTCGGAATCAGACGAGCGAGCCAGGAGAGAGAGTGGAAATGGCGGCA 1620
DB 1561 CGCGCCCTCGGAATCAGACGAGCGAGCCAGGAGAGAGAGTGGAAATGGCGGCA 1620
QY 1621 CTAACGTGATCAAACTAACCGGAAATCACAATCTTAAGGGTTCCTTTCGAAAACGTG 1680
DB 1621 CTAACGTGATCAAACTAACCGGAAATCACAATCTTAAGGGTTCCTTTCGAAAACGTG 1680
QY 1681 TCGTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAAGCTGGGAGATTCGCTCATCGGCG 1740
DB 1681 TCGTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAAGCTGGGAGATTCGCTCATCGGCG 1740
QY 1741 TTGCTGAGAGCGAGCGCGCGGAGCAAGAGCTGAGCTCCAGCTCGAGCTTCTTAAA 1800
DB 1741 TTGCTGAGAGCGAGCGCGCGGAGCAAGAGCTGAGCTCCAGCTCGAGCTTCTTAAA 1800
QY 1801 TCATCTTCAAGTACATGGATTTTATTTGCTCTTTGCTTTGCTTAAAGTGTACTAT 1860
DB 1801 TCATCTTCAAGTACATGGATTTTATTTGCTCTTTGCTTTGCTTAAAGTGTACTAT 1860
QY 1861 GCGATGAAGAAATACCAAGTATAGCAAGCTGAGGTGTGTAGTACTACTAGAAGTCA 1920
DB 1861 GCGATGAAGAAATACCAAGTATAGCAAGCTGAGGTGTGTAGTACTACTAGAAGTCA 1920
QY 1921 GTCACGTTCTTCTGTAAGAAATGTTAACTGTTAATTAAGCAGTATGTTTCAGTATC 1980
DB 1921 GTCACGTTCTTCTGTAAGAAATGTTAACTGTTAATTAAGCAGTATGTTTCAGTATC 1980
QY 1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
DB 1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017

RESULT 2

US-09-291-922-25
; Sequence 25, Application US/09291922
; Patent No. 6383775
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-25

Query Match 54.3%; Score 1094.6; DB 4; Length 2089;
Best Local Similarity 79.1%; Pred. No. 8.7e-206;
Matches 1343; Conservative 1; Mismatches 335; Indels 18; Gaps 3;

QY 84 GCACCAACCGAGATGCTTCCCGCGGTTCGCGGAGCGGTTCGCGGAGAGAGGCG 143
DB 65 GGCTCGTCAGCATGACCGCGCGCTCCCGCGCGCGTTCGAGCCCAAGAGAGGCG 124
QY 144 AAGTCGCGTTCGCTTCGCTTCGCGCATCTTCGCTCCATGACCTCCATCTCTCGC 203

			125 AACGTGAGTTGCCTTTCGGCTCGGCATCCTCGCCCTCAATGACCTCCATCCTCTCGGC	184
Db				
Qy			204 TACGATATCGGGTGATGAGCGGGCGCTCGTGTACATCAAGAAGACTTCAACATCAGT	263
Db			185 TAGACATCGGCGTGATGAGCGGAGCGTCTGCTGTACATCAGAGGATCTGAAGATCAAC	244
Qy			264 GACGGGAAGTGGAGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC	323
Db			245 GACACCAGCTGGAGTTCCTCATGGGCATCCTAACGTGTACTCGCTCATTTGGTCTCCTC	304
Qy			324 CGSCGGGGCGGACGTCGGACTGATCGGCGGGGTACACCATCGTGTTTCGCCGCGCTC	383
Db			305 CGSGCGGGCGGACCTCGACTGGATCGGCGGGCGCTTACCATCTGTTTCGCCGCGCTC	364
Qy			384 ATATTCTTCGCGGGGGTTCCTCATGGGTTTCGCTCAACTACGCCATGCTCATGTTTC	443
Db			365 ATTCTTCTCGCGGGCGCCTCATCATGGGTTTCTCCGTCAACTACGCATGCTCATGTTTC	424
Qy			444 GGCGCTTCGTGGCGGCATCGGCGTGGGTACGCGTCTCATCTTCCTCCGAGGTGTTTCATC	503
Db			425 GGGCGCTTCGTGGCGGCATCGGCGTGGGTACGCTCTCATGTATCGCGCGCTGAACAGC	484
Qy			504 GCGAGGTGTCGCGGGCGTGGCGCTGGGTTCTCTGACGTGTTTCCCAGAGTGTTTCATC	563
Db			485 GGGAGGTGCCCCCGGTCTGCGCGTGGGTTCTCACATCTTCCCGAGGTGTTTCATC	544
Qy			564 AACTTCGGCATCTGCTCGGTACGTTCTGNACTATGCTTTCCTCCGCTTCCCGCTGAAC	623
Db			545 AACTTCGGCATCTCTCCGATATGTCTCAACTTCGCTTCGCGCGCTCTCCCTCCGC	604
Qy			624 CTCGGTGGCGCATCATGCTCGGCATCGGCGCGCGCGCTCCGTGCTGCTCGCGCTCATG	683
Db			605 CTCGGTGGCGATTATGCTCGGCATAGCGCGGTGCOGTCTGCTCTGCTCGGTTTCATG	664
Qy			684 GTGCTCGGCATCGCGAGTCCGCGGTGTGTCATGAAGGAAGCCCTCGCGGAGCCG	743
Db			665 GTGCTCGGCATGCCCGAGTCTCCCGGTGGTCTGCTCATGAAGGCCCTCTCGCGGAGCCG	724
Qy			744 AAGTGTGTGTGGAGAAGACTCCGACACGCGGAGGAGCGCGGAGCGCTGGCGGAC	803
Db			725 AAGTTGTGTTGCAAGAGTCCGACACGCCGGAAGAGCGCGGAGCGATCGCCGAC	784
Qy			804 ATCAAGGCCCGCGCGCATCTCTGAGGAGTTCGACGGCGACGTGTGTGACCGTCCCC--	860
Db			785 ATTAAGACTCGCGCGCATCTCTCGGCGCTCGACGGCGAGTGTCTCCCGTGCCCAA	844
Qy			861 AAGAGGGGAGCGAACAAGAGCGGTGTGSAAGGAGTCTATCTTCTCCCGACCCCG	920
Db			845 AACAAAGGAAGCAGGAGGAAGCGGTTTTGAAGGACCTCATCTGTCAACGACCAT	904
Qy			921 GCCATCGGGGCATCTCTGCTGTCCGGATCGGCATCCACTTCTTCCAGCATGTGTGGGC	980
Db			905 GCCATCGCCACATCTCTCATCGGGAAATCGGCATCCACTTCTTCCAGAGTCTTCGGC	964
Qy			981 ATTCATCTCGTCTGTTCTACAGCCCTCTGTGTTCAAGAGCCCGGATTAAGAGCAC	1040
Db			965 ATGACCCCGTGTGCTCTACAGCCCTAGTTTTTCAAGAGCGCGGCATCAAGGGGAC	1024
Qy			1041 AAACATCTTGGGCACCATTTGCGCGTGTGTTGTACCAAGAGGCTTTTCATCTGTTG	1100
Db			1025 AGCGCTCTCGCGSACCAACGTTGGCGGTGGGGCGACCAATACGTTCTTCACTCTGTTG	1084
Qy			1101 GCAGCTTTCTTATGACGGGTTCGGCGCGGCGCTGTGTTGTGGGAGCAGCGGGCGG	1160
Db			1085 GCCACTTCTCTCTGACCGCATCCGCGCGCGCGCTGTGTTGTGACACAGCGGGCGG	1144
Qy			1161 ATATCTCTCTCCCTCATCGGCTCGGCGCGGGGTCAAGGTGCTCGCGAGCACCCGAC	1220
Db			1145 ATGCTGTCTCTTAGTGGGCTTCGACGGGGTTCACCGTATCAAGCCGACCCGAC	1204
Qy			1221 GCCAAGATACCTTGGGCGCATCGGCTAAGCATCGCCCTCCACCTTCGCTACGTCGCTTC	1280

Db	1205	GAGAAGATCACTGGGCGATCGTCTGTGCATCTTCTTGCAATCAATGGGCTACGTGGCGCTTC	1264
Qy	1281	TTCTCCATCGGCTTGGCGCCCATCACTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG	1340
Db	1265	TTCTCCATCGGCTTGGCGCCCATCACTGGGTGTACAGCTCGGAGATCTTCCCGCTGCAC	1324
Qy	1341	GTGGCGGCTGGGCTGCTGCTCGGCGTGGCGCCAAACCGGTCAACGAGCGGCGTCA	1400
Db	1325	GTGGCGCGTGGGCTGCTGCTGGGCGGTGGCGGTCAACCGCTGACGACGCGGCGTGATC	1384
Qy	1401	TCCATGACCTTCCCTGTGCTGTCCAAAGCCCATCAACCATCGGCGGCAGCTTCTTCTCTAC	1460
Db	1385	TCCATGACCTTCAATTCGCTGTCCAAAGCCCATGACCATCGGCGCGCTTCTTCTCTTC	1444
Qy	1461	TCGGCATTCGGCGGCTGCGCTGGGTGTTCTTACACCTACCTCCCGGAGACCGCGGC	1520
Db	1445	GCGGCGATCGGCTCAATTCGATGGGTGTTCTTCTTCGCTCACTTCGCGGAGACCGCGGC	1504
Qy	1521	CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCGCGCGCTCGGAATCAGAC	1580
Db	1505	CGCACGCTGGAGACATGAGCTCGCTGTTTCGGCAACACGGCCACGACACAGCAGGCGGC	1564
Qy	1581	GAGCCAGCCAAAGAGAGAAAGAGTGGAAATGGCCGCCACTAACTGATCAAACTAACCG	1640
Db	1565	GCGAAAGCCAGCAGCAGCGCGGGGAGAAAGGTGGAATGGCGCGCACCAACTGACCG	1624
Qy	1641	CAAAATCACCAATCCTTAAGGGTTTTCTTCGAAAACGTTGCTGTACTGCTAGCTAGC	1700
Db	1625	CAAGTTGGCAGATCGCGATCGGAAGACTTGGCGTGTATCCGCTCTGGGTAGCTAGCTGCC	1684
Qy	1701	A-----AGTAGTAGCAGCAACGTCGGGAAGATTTCGTGATCCGGC----GTTGCT	1748
Db	1685	ACAAGGCCACATAGATGACGAAGTAGCGTGGGAAGATTGCTGATCCGCGGAGCTGCC	1744
Qy	1746	GGAGAGCGACGGCCGGC	1762
Db	1745	GGAGGCGACGGCAAGC	1761

RESULT 3

US-09-291-922-19

; Sequence 19, Application US/09291922

; Patent No. 6383776

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: BB-1163

; CURRENT APPLICATION NUMBER: US/09/291,922

; CURRENT FILING DATE: 1999-04-14

; EARLIER APPLICATION NUMBER: 60/083,044

; EARLIER FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 19

; LENGTH: 1914

; TYPE: DNA

; ORGANISM: Zea mays

US-09-291-922-19

Qy	95	GATGGCTTCGCGCGCTGCGGAGGCGCTCGCGCCGGAAGAGCGCAACGTCGCGTT	154
Db	55	GATGGCTTCGACGAGCTGCGAAAGGCGCTCGAGCCAGGAGGCGCAACGTCAGTA	114
Qy	155	CGCTTCGCTTCGCGCATCTTCGCTTCATGACCTCCATCTCTCGGTACGATATCGG	214
Db	115	TGCTTCATATGTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG	174

Query March 49.8%; Score 1005; DB 4; Length 1914;

Best Local Similarity 79.2%; Pred. No. 3.2e-188;

Matches 1220; Conservative 1; Mismatches 311; Indels 9; Gaps 2;

RESULT 3

```

US-09-291-922-19
; Sequence 19, Application US/09291922-19
;
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Fingsky, Scott
; TITLE OF INVENTION: Plant Sugar Tra
; FILE REFERENCE: BB-1153
; CURRENT APPLICATION NUMBER: US/09/22-19
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-09-291-922-19

```

	Query Match	49.8%	Score 1005	DB 4	Length 1914
	Best Local Similarity	79.2%	Fred. No. 3.2e-386		
	Matches 1220	Conservative	1	Mismatches 311	Indels 9; Gaps 2
Qy	95	GATGGTTCCGCGCGTGC	CGGAGGCGCTGCGCGCGAA	GAGAAAGGCGCAACGTC	CGGTT 154
Db	55	GATGGTTCCGACGAGCTCG	CAAAAGCGCTGACGCCAG	GAAAGGCGCAACGTC	CAAGTA 114
Qy	155	CGCTTCGCGCTGGCGCAT	CTCGCCTCCATGACCTCC	ATCCTCTCGGCTACGATATCG	214
Db	115	TGCTTCATATGTGCATCT	TGCGCTCCATGGCTCTGT	CATCTCTGGCTATGACATTTG	174

Db 428 TCCTTGCTCATGGGCTTCGCGGTCAACTACGCCATGCTCATGTGTCGGGCGCTTCGTCAAC 487
QY 459 GGCATCGCGGGCTACGGCTCATGATCGCGCGGTGTACACCGCGGAGGTGTGCGCG 518
Db 488 GGAATCGCGGTGGCTTACGCCATCATGTCGCGCAAGTGTACAGCCCGAGGTGTCCCG 547
QY 519 CGCTCGCGGTGGCTTCCTGACGTGCTTCGCGGAGGTGTTCATCAACTTCGGCATCCTG 578
Db 548 CGCTCGCGCGGTTCCTCACTGCTTCACCGAGGTGTTCATCAATGTGGGCATCCTC 607
QY 579 CTCGGGTACGTCGAACTATGCTTTCCTCCGCTTGGCGTGAACCTCGGTCGCGCATC 638
Db 608 CTTGGCTACGTCCTCAACTACGCTTCGCGCGCTCCCGTCCACCTCAGCTGCGCGTC 667
QY 639 ATGCTCGGCATCGCGCGCGCTTCGCTGCTGCTCGCGCTCATGTCGTCTCGGCATCGCG 698
Db 668 ATGCTCGGCATCGCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727
QY 699 GAGTCGCGCGGTGGCTGTGATGAAGGAGCGCTTCGCGGACGCAAGGTGTGCTGGAG 758
Db 728 GAGTCCTCGCTGGCTGTGATGAAGGAGCGCTTCGCGGACGCAAGGTGTGCTGGAG 787
QY 759 AAGACCTCCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGACGCAAGGTGTGCTGGAG 818
Db 788 AAGACCTCCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGACGCAAGGTGTGCTGGAG 847
QY 819 GGCATCCCTGAGGAGTCGACGCGGAGCGCTTCGCGGACGCAAGGTGTGCTGGAG 878
Db 848 GGCATCCCTGAGGAGTCGACGCGGAGCGCTTCGCGGACGCAAGGTGTGCTGGAG 907
QY 879 GAGAGCGGCTGGAAGGAGTCATCTGTCGCGGAGCGCTTCGCGGACGCAAGGTGTGCTGGAG 938
Db 908 GAGAGCGGCTGGAAGGAGTCATCTGTCGCGGAGCGCTTCGCGGACGCAAGGTGTGCTGGAG 967
QY 939 CTGTCCGCGGATCGGCATCCATCTTTCAGCATCGCTTGGGATTCATCTCGCTGCTTC 998
Db 968 CTCGGCGCTCGGATCCATCTTTCAGCATCGCTTGGGATTCATCTCGCTGCTTC 1027
QY 999 TACAGCCTCTCTGTGTTCAAGAGCGCGGATTAAGAAAGCAACATCTTGGGAGCC 1058
Db 1028 TATAGCCACGCGTGTTCAGAGCGCGGATTAAGAAAGCAACATCTTGGGAGCC 1087
QY 1059 ACTTGGCGCTGCGTGTACCAAGAGGCTTTCATCTTGTGGGAGTTCCTTCATCGAC 1118
Db 1088 ACATGCGCATGGGGTCAAGAAAGCGCTTTCATCTTGTGGGAGTTCCTTCATCGAC 1147
QY 1119 GGCCTCGCGCGCGCGCTGTGTCGCGGAGCAAGCGCGGATTAAGAAAGCAACATCTTGGGAGTTCCTTCATCGAC 1178
Db 1148 GGCCTCGCGCGCGCGCTGTGTCGCGGAGCAAGCGCGGATTAAGAAAGCAACATCTTGGGAGTTCCTTCATCGAC 1207
QY 1179 GGCCTCGCGCGCGCGCTGTGTCGCGGAGCAAGCGCGGATTAAGAAAGCAACATCTTGGGAGTTCCTTCATCGAC 1238
Db 1208 GGCCTCGCGCGCGCGCTGTGTCGCGGAGCAAGCGCGGATTAAGAAAGCAACATCTTGGGAGTTCCTTCATCGAC 1267
QY 1239 ATCGGCTTAAGCATCGCTTCACGCTTCAGCTGCTGCTTCTTCATCGGCGCTTGGC 1298
Db 1268 ATCGGCTTAAGCATCGCTTCACGCTTCAGCTGCTGCTTCTTCATCGGCGCTTGGC 1327
QY 1299 CCAATCAGTGGGTGTACAGTTCGAGATCTTCGCGTTCAGGTTCGCGGCGCTGGGTGC 1358
Db 1328 CCAATCAGTGGGTGTACAGTTCGAGATCTTCGCGTTCAGGTTCGCGGCGCTGGGTGC 1387
QY 1359 TCGCTCGCGCTCGCGCGCTTCACGCTTCAGCTGCTGCTTCTTCATCGGCGCTTGGC 1418
Db 1388 TCGCTCGCGCTCGCGCGCTTCACGCTTCAGCTGCTGCTTCTTCATCGGCGCTTGGC 1447
QY 1419 TGTCCAAAGGCATCAACATCGCGGAGCTTCTTCCTTACTCCGCGATCGCGCGCTC 1478
Db 1448 TGTCCAAAGGCATCAACATCGCGGAGCTTCTTCCTTACTCCGCGATCGCGCGCTC 1507
QY 1479 GCCTGGGTGTTCTTACTACATCTCCTCCGAGACCGCGCGGATCGCTGGAGGATG 1538

Db 1508 GGATGGATTTCTTCTTACCTTCAATTCGCGAGACGCTGGCTCGCGCTCGAGGAGATA 1567
QY 1539 AGCAAGCTTCTCGCGACACGCGCTCGGAATCAGAGCCAGCCAGGAGAGAG 1598
Db 1568 GGAAGCTTTTCGCGATGACGACGCGCTCGGAAGCCAGACACCGCCAGAGAC 1627
QY 1599 AAGAAGTGAAGTGGCGCCCACTACTGATCAA 1632
Db 1628 AAGCGAAAGTAGGGGAGATGAAGTAGTGAGCTA 1661

RESULT 5

US-09-291-922-23
; Sequence 23, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-23

Query Match 28.5%; Score 575.2; DB 4; Length 1853;
Best Local Similarity 63.6%; Pred. No. 5.9e-104;
Matches 909; Conservative 2; Mismatches 510; Indels 9; Gaps 2;

QY 109 CGCTGCGGAGCGCTGCGCGCGGAGAGAGGGAACGTCGCTTCGCTTCGCTCGCG 168
Db 210 CACTTCAGGATTCGATCTCCAAAGAGCGGAAAGAACAGTATGCTTTGCTTGTG 269
QY 169 CCATCTCCCTCCATGACCTCCTCCTCGGTACGATATCGGGGTGATGAGCGGG 228
Db 270 CTATGCTGCTCCATGACTTCCATCTTCTGTTATGATATGAGTATGAGTGGAG 329
QY 229 CGTCTGTACATCAAGAGGACTTCAACATCAGTGAACGGAAGTGGAGTTCATCG 288
Db 330 CAGCCATATACATAAAAGGAGCTGAAAGTCTCGAGCAGCAATCGAGATCTCTCG 389
QY 289 GCATCTGAACCTTACTGCTCATCGGCTCCTTCGCGCGGGGCGAGCTCGGACTGGA 348
Db 390 GAATCATCACTTACTCTCTGATAGGCTCATGTCGCGCGAGAACTTCGACTGGA 449
QY 349 TCGCGCGCGGTACACCATCGTGTTCGCGCGCTCATATTTTCGCGGGGGGTTCCTCA 408
Db 450 TAGTCCCGCTTACAGATTTGTTTCGCGCGCACCATCTTCTTGTGCGAGCACTTCTCA 509
QY 409 TGGGTTCGCGCTCAACTACGCTCATGTCGCGCGCTTCGTCGCGCGGCTTCGCGCG 468
Db 510 TGGGTTCGCGCGCTCAACTACGCTCATGTCGCGCGCTTCGTCGCGCGGCTTCGCGCG 569
QY 469 TGGGTTCGCGCGCTCAACTACGCTCATGTCGCGCGCTTCGTCGCGCGGCTTCGCGCG 528
Db 570 TGGGTTCGCGCGCTCAACTACGCTCATGTCGCGCGCTTCGTCGCGCGGCTTCGCGCG 629
QY 529 GTGGCTTCGCGCTCAACTACGCTCATGTCGCGCGCTTCGTCGCGCGGCTTCGCGCG 588
Db 630 GTGGCTTCGCGCTCAACTACGCTCATGTCGCGCGCTTCGTCGCGCGGCTTCGCGCG 689
QY 589 TCTCGAACTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 648
Db 690 TATCAAACTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 749

```
QY 649 TCGGCGCGCGCGCTCGTGTCTCGGCTCATGTTCTCGGATCGCGAGTCGCGC 708
Db 750 TTGGTGCATACCTTCGGTACTCTTAACAGTAGAGTGTGGCATGCCGAGTCCCA 809
QY 709 GGTGCTGTCTATGAAGGAGCGCTTCGCGGACGCAAGTGTGTCTGGAGAACCTTCG 768
Db 810 GGTGCTGTGTATGAGGCGTGTGTTGGGAGAGCAAGAAAGTGTCTTAACAAACCTCAG 869
QY 769 ACACGCGGAGGAGCGCGGAGCGCTTGGCGACATCAAGCGCGCGCGCATCTCTG 828
Db 870 ACAGCAAGAGAGCGCCCACTAAGCTTAGCGGAATCAAAACAGCGCGAGATCCCG 929
QY 829 AGGAGCTCGACCGCGACGTGTGTGACCTGCCCAAGAGAGGAGCGGAAACGAGAAGCGG 888
Db 930 AGAGTTGCAACGACGACGTGTGTTAGGTAATAAACAAGCAACGGTGAAG- ----GTG 983
QY 889 TGTGAAGAGCTCATCTGTCTCCCGACCGCGGCATCGCGGATCTGTCTCGGGA 948
Db 984 TATGGAAGAGCTCTCTCTATCCACGCGCGCAATTCGTACATCGTAAATCGCTGCC 1043
QY 949 TCGGATCCACTTCTTCAGATGCTTGGGCAATTCACCTCGGTGTCTTACAGCGCTC 1008
Db 1044 TTGGTATTCACTTCTTCCACAGCGCTCGGCGGTAGACGCGTGTGTTGTACAGCCCA 1103
QY 1009 TCGTGTCAAGAGCGCGGATTAAGCAAGCAAAACATCTTGGGACACCACTTGGCGGT 1068
Db 1104 GGATCTTCCAAAGCTGGGATTAACAACGACACGATTAAGCTTCTTGAACCGTGGCG 1163
QY 1069 TCGGTGTACCAAGAGGCTTTTCTCTTGTGGGACTTCTTCAAGCGCGTTCGGC 1128
Db 1164 TTGGATTCGTGAACCGTGTCTTCTTGGCGGTACGTTTACGTTGAGCGCGTGGGTC 1223
QY 1129 GCGCGCGCGTGTGTGGGAGCAGCGCGGAGTAATCTCTCCCTCATCGCGCTCGCG 1188
Db 1224 GTGTCGCGTGTGTTATGTTGTAGTGTGCGCGCATGGTGTCTCTACGCTTGGCA 1283
QY 1189 CCGGCTCACGCTGTCGCGCAGACCGCGGAGCCAGACATACCTTGGCGCATCGGCTAA 1248
Db 1284 TCAGCTTCACTGTTATG- - -ATCATTGGAAGAAATTAATGTGGCGCTTGGATCGA 1340
QY 1249 GCATCGCTTCCACCTCGCTACGTCGCTTCTTCTCCATCGGCTTGGCGCGCATCAGT 1308
Db 1341 GCATAGCATGGTGTGGTGTAGTGGCGACGTTCTCCATCGGTGCGGTCCTCAATCAGT 1400
QY 1309 GGGGTACAGTTCGAGATCTTCGCGTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGCG 1368
Db 1401 GGGTCTATAGTTCGAGATCTTCGCTTGGGCTGCGCGCGCARGGTGCGCGCGGAG 1460
QY 1369 TCGCGGCAACCGCTCACGAGCGGTCATCTCCATCAGCTTCTGTGCTGTCCAAG 1428
Db 1461 TTGCGGTGAATAGACACTAGCGCGTGTCTCAATGACTTCTTCTCCCTCACTAG 1520
QY 1429 CCATCACCATCGCGGAGCTTCTCTCTACTCGGATCGCGCGCTCGCGCTGGGTGT 1488
Db 1521 CCATCAGTATGGTGGAGCTTCTCTCTTATGTCGATCTGCTACTGTTGGGTGGATAT 1580
QY 1489 TCTTCTACACTTACCTCCGAGACCGCGCGGAGCCTGAGGAGATG 1538
Db 1581 TCTTTTACCGCTTGTGCTGAGACCGCGGAAACCTTCGAGACATG 1630
```

RESULT 6

```
US-09-291-922-17
; Sequence 17, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
```

```
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (149)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (271)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (304)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (334)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (357)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (476)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (599)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (602)
; LOCATION: (602)
; US-09-291-922-17

Query Match 17.0%; Score 343.6; DB 4; Length 615;
Best Local Similarity 85.6%; Pred. No. 1.1e-58;
Matches 423; Conservative 1; Mismatches 66; Indels 4; Gaps 4;

QY 84 GCACCGCGGAGATGCTTCCCGCGCTGCGCGAGCGCTGCGCGGAGAGAGGCG 143
Db 65 GAACCGTCAACGATGCGCTTCCGCTGCGCGGCGCATCGAGCGCGGAGAGGCG 124
QY 144 AACGTCGCGTTCGCTTCCGCTGCGCATCTTCCGCTTCCATGACCTTCCATCTCTCGGC 203
Db 125 AACGTCAGTTTCCGCTTCCGCTTCCGCTTCCGCTTCCATGACCTTCCATCTCTCGGC 184
QY 204 TAGCATATCGGGTGTATGAGCGGCGCTGTGTATCATCAAGAGAGCTTCAACATCAGT 263
Db 185 TATGATATCGGATGATGAGCGGCGCTGTTGTATCATCAAGAGAGCTTCAACATCAGC 244
QY 264 GACGGAAGGTGGAGTTCCTCATGGCATCTGAACTTCTACTCTGCTCATCGGCTCTTC 323
Db 245 GAGTGAAGGTGGAGTTCCTCATGGCATCTGAACTTCTACTCTGCTCATCGGCTCTTC 304
QY 324 GCGCGCGGCGGAGCTGCGGATGAGTGGCGCGGCTACCATCTGTGTTCGCGCGCTC 383
Db 305 GCGCGAGGCGGAGCTGCGGATGAGTGGCGCGGCTACCATCTGTGTTCGCGCGCTC 363
QY 384 ATATTTCTTCGCGGCGGCTTCTCATGGGTTTCGCGCTCACTAGCCATGCTCATGTT 443
Db 364 ATCTTCTTCGCGGCGGCTTCTCATGGGTTTCGCGCTCACTAGCCATGCTCATGTT 422
QY 444 GCGCGCTTCTGCGCGGCGCATCGGCTGAGTGGGCTACGCGCTCATGATCGCGCGGTACACC 503
Db 423 GCGCGCTTCTGCGCGGCGCATCGGCTGAGTGGGCTACGCGCTCATGATCGAACCGTACAG 482
QY 504 GCGGAGGTGCGCGGCGGTGCGGCGGTGCTTCTGAGCGCTTCCGCGAGGTGTTCATC 563
Db 483 GCGGAAGTGTG-CCGCGCATCGGCGCGGCTTCTGAGCGTCTTCCGCGAGGTGTTCATC 541
QY 564 AACTTCGCGCATCTCT 577
```

Db 542 -ACTCGGCATCCT 554
|||||
527 GGTGGCTTCCTGACGCTGTTCCCGAGGTGTTTCATCAACTTCGCGCATCTCTGCTCGGTA 586
568 CCGGGCATGCTCAACATCGGCTTCAGCTGATGATACCATCGGCATCTCTCGCGCGGA 727
587 CGTCTCGAATATGCTTTCTCCGCTTCCGCTGACCTCGGTTGCGCATCATGCTCG 646
728 GCTCATCAACTACGGACCAACAAGATCAAGCGGGGTACGGGTGCGGTGAGCCTGCG 787
647 CATCGCGCGCGCGCTCCGCTGCTGCTCGCTCATGTTGCTCGCATCCCGAGTTCGC 706
788 GCTGGCGCGGTGCGCGCGCATCATCACTCGGCTCTCTCTCTCTCTCTCTCTCTCTCT 847
707 GCGTGGCTGCTGATGAAGGAGCGCTCGGACCGCAAGTGGTGGTCTGCTGGAAGACTTC 766
848 CAACTCGCTGAGCGGGGCAACCGGAGGAGCAACCGCATGCTCTCGCGCATCG 907
767 CGACACGCGGAGGAGCGCGCGAGCGCTGCGCGCATCAAGGCGCGCGCGCATCGC 826
908 CGGACGAGACATCGCGGAGGAGTACGCGACCTGTTGGCGGCGAGGAGGCGCG 967
827 TGAGGAGCTCGACGCGGAGCGTGTGACCTGCCAAGAGAGGAGCGGAAACGAGAGCG 886
968 CCAGGTGCGCCACCC-----982
887 GGTGTGAAGGAGCTCATCTGTCGCGACCGCGCATGCGCGCATCTCTGCTCGCG 946
983 ----GTGGCGAATCTCTGCGCGCGGTACCGCGCGAGCTCACCAT-----GGC 1030
947 GATCGGCATCCACTTTCTCCAGCATCGTTGGGCAATCACTCGTGTCTTCTACAGCCC 1006
1031 CGTGGCATCCCTCTCTCCAGCAGCTCAAGGGATCAACGTCATCATCTTCTACGCGCC 1090
1007 TCTCGTGTTCAGAGCCCGGATTAAGCAACGACAAACATTTCTGGGCAACACTTGGCC 1066
1091 CGTGTGTTTCAGACGCTGGGATCAAGAAAGAGCTTCTCTCATGTCCTCGTCATCAC 1150
1067 GTTGGGTGTCACCAAGAGCTTTTTCATCTTGTGGGAGCTTTCTTCATCGAGCGGTTCGG 1126
1151 G---GGCTGTCGACGCTTTCGCCACGCTGTCGTCGTCACGTCGTCACGCGCTCGG 1207
1127 GCGGCGCGCGCTGTTGCTGGGAGCAGCGGCGGGATATCTCTCTCTCATCGGCTCGG 1186
1208 CCGCGCAAGCTGTTCTCTCCAGGCGGCGGAGATGATGTCGTCGCGAGCTCATCTGGG 1267
1187 CGCGGGCTCACGCTGCTCGGCGAGCACCGCGACGC-----CAAGATACCTTGGGCCAT 1240
1268 CAGGCTCATCGCGCAAGTTCGGACACAGCGGACGCGGGGACATCGCAAGGCTACGC 1327
1241 CGGCTTAAGCATCGCTCCACCTCGCCTACGCTCGCTTCTCTCTCTCTCTCTCTCTCTCT 1300
1328 CGCGGTGCTGCTGCTCTCTCATCTGCGCTACGTCGCGGCTGCTGCTGCTGCTGCTGCT 1387
1301 CATCAGTGGGTGACAGCTCGAGATCTTCGCTCCAGGTGCGGCGCTGGGCTGCTGCT 1360
1388 CTGCGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1447
1361 GTCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1420
1448 CATCAAGCTCTCGTCAACATGTTCTTCACTTTCGATCGCGAGGCTTCTCTCACCAT 1507
1421 GTTCAAGGCGCATCAGATCGGCGGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1480
1508 GCTCTGCCACTTCAAGTT---CGGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1564
1481 CTGGGTGTTTCTTCTACCTACTCTCCGAGACCGCGCGCGGCGGCTGGAGAGATG 1538
1565 GACCGTCTTATCGCTTCT 1622

RESULT 8
US-09-679-686B-11
; Sequence 11, Application US/09679686B
; Patent No. 6624343

Query Match 9.4%; Score 188.8; DB 4; Length 1675;
Best Local Similarity 50.1%; Pred. No. 3.3e-28;
Matches 660; Conservative 1; Mismatches 588; Indels 69; Gaps 5;
5;
QY 227 GCGTGGCTGCTGATCAAGAGGACTTCAACATCAGTGACGGGAAGTGGAGTTCAT 286
Db 368 GCGCACCTGACGAGCGGACTACTCAAGTACGACCAACAGTGTGCTGCTCTTTCAC 427
QY 287 GCGCATACTGAACCTCTACTCGCTCATCGGCTCTCTCGCGCGGCGGAGCTCGGACTG 346
Db 428 CTCGTGCTCTACTTCGCGGCTCTGCTCACTTCGCGGCTCTCTACGTGACCAAGCG 487
QY 347 GATCGCGCGGCTGACCATCTGTTGCGCGCGCTCATATCTTCGCGGCGGCTGCTCT 406
Db 488 CCACGCGCGGCGCGGAGCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
QY 407 CATGGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466
Db 548 CACGCGCGCGGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607
QY 467 CGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
Db 608 CGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667

GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: B01160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1776
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (66)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (219)
; OTHER INFORMATION: n = a, c, g or t
US-09-679-686B-11

Query Match 8.3%; Score 168.4; DB 4; Length 1776;
Best Local Similarity 48.8%; Pred. No. 3.4e-24;
Matches .632; Conservative 1; Mismatches 597; Indels 66; Gaps 4;

QY 243 AAGAGGAGCTTCAACATCAGTACGAGGAGGTGGAGGTCTCTCATGGGCATCTCAACTC 302
DB AACAATCTAGTGAAGTACAGACAGAGGCTCTCGGGTTTCACTCTCTCTACTCTC 363
QY 303 TACTCGCTCATGGCTCTTTCGCGCGGGGGAGAGTTCGAGTCTGGATCGCGCGGTAC 362
DB GCGGCGCTCGTCTCTCTCTGCGGCTCGCGGCTGACGAGGAACTACGCGCGCGCGCC 423
QY 363 ACCATCGTTCGCGCGCTCATATCTTCGCGGGSSETTCTCATGCGGTTTCGCGCTC 422
DB AGCATCGTTCGCGCGGCTCAGCTTCTCTCGCGGCGCACGCTCAACGCTCTCGCGCTT 483
QY 423 AACTACGCGCATCTCATGTTTCGCGCGCTTCGTCGCGGCGCATCGGCTGGGTACGCGCTC 482
DB AACCTCGTCATGCTCATCTCTCGCGCGCATCTCTGCTCGCGCTCGGAATCCGCTTCGCGAAC 543
QY 483 ATGATCGCGCGGTGTACACCGCGAGGTGTGCGCGGCTCGCGGCTGCGGCTTCTCTGACG 542
DB CAGGCGGTGCGGCTGTACTCTCTCGGAATGCGCGGCGCGCACCTTCGCGCGCGCTGAAC 603
QY 543 TCGTTCGCGGAGTGTTCATCAACTTCGCGCATCTCTGCTCGGGTACGCTCTCGAACTATGCT 602
DB ATGATGTTCCAGTCTGGCGACGAGCTGGGCATCTTCACGGCGAATGATCACTACGCG 663
QY 603 TTCTCCCGTTGCGCTGAACCTCGGGTGGGCATCATGCTCGGCATCGCGCGCGCGCG 662
DB ACCGACACATCAGGCGG---TGGGGGTGGGGGTCTCTCGTGGGGCTCGCGCGCGCGCG 720
QY 663 TCCGTGCTGCTCGGCTCATGTTCTCGGCATCGCGGAGTTCGCGGCTGGCTGCTGATG 722
DB GCGCTGCTGATGACGTTGCGGGGTGCTCTCTGCGGAGAGCGCCACAGCTGATCGAG 780
QY 723 AAGGACGCTTCGCGACGCCAAGGTGTGTGTGAGAGAGCTTCGACACGCGGAGAG 782
DB CCGCGGCGGTGAGAGGCGCGCGCTGTGTGAGCGCATCTCGGGGACACGCGCGAGGTG 840
QY 783 GCGCGGAGCGCTGCGCGCATCAAGGCGCGCGCGCTCCCTGAGGAGCTCGAGCGC 842
DB GACGCGGAGTTTACGAGCATGCGGAGGC----- 869

QY 843 GACGTGTTGACCGTCCCCAAGAGAGAGGAGCGGAAACGAGAGCGGTGTGTGAAGAGTTC 902
DB -----GAGCGAGCTCGCAACTCCATCGAGCACCGGTTC 903
QY 903 ATCTGTCCCGACACCGCGCATCTCTGCTGTCGGGATCGGATCGGATCTCACTTC 962
DB CGCAACATCTGAGCGCGCAACCGCGCGAGCTGTTGATCGCGGTGTGATCGCGCG 963
QY 963 TTCCAGCATGCTGGGATTCACCTCCGTCGCTTCTACAGCCCTCTCTGTTTCAAGAGC 1022
DB TTCCAGATCTGACGGGCATCACTCCATCTCTTCTACGCGCGCTGCTGTTCAGAGC 1023
QY 1023 CCGGATTAAACGAGACAAACATCTTTGGGACCACTTTGGGACCACTTTGGGTCACCAAG 1082
DB ATGGGCTTCGG---CGGACGCGTCTACTCTCTAGTCTCACGCGCGCGCTCTC 1080
QY 1083 AGCTTTTTCATCTTGTGGGACTTCTTCATCGAGGCGTTCGGGCGCGCGGCTGTGTG 1142
DB TTCTCTCCACCATCACTCTCCATCTCCACCGTCCGCGCTCGCGCGCGCGAGCTCTCT 1140
QY 1143 CTGGGACGACGCGCGGGGATAATCTCTCTCTCATCGGCTCGGCGCGCGGCTCACCGTC 1202
DB ATCAGCGCGGCATCCAAATGATCATCTCGCAGGTGATAGTGGCGTGTCTTGGGGTG 1200
QY 1203 GTGGCGCAGACCGCGACCCAGATACCTTGGGCGCATCGGCTTAAGCATCGCTCCACC 1262
DB AAGTTCGGGACGACAAAGAGCTGACGAGAGCTACTCGATCGCGGTGTGTGTGATC 1260
QY 1263 CTGCGCTACGTCGCTTCTTCTCCATCGGCTTGGCGCGCATCACGTGGGTGTACAGCTCG 1322
DB TGCTGTTTGTGCTGCGTTCGGGTGCTGCGGGGCGCTCGGGGTGGAGCTGCGGAGC 1320
QY 1323 GAGATCTTCCGCTCAGGTGCGCGCTGGGCTGCTGCTGCGGCTCGCGCGCGCAACCGC 1382
DB GAGATCTTCCGCTGAGACGAGGTGCGCGGCGGAGAGCATCACGTGGCGGTGAACCTC 1380
QY 1383 GTACACGAGCGGTGATCTCCATGACCTTCTGCTGCTTCCAGGCGCATCACCATCGGC 1442
DB TTCTTCACTTCTGATCGCGAGGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1443 GGAGCTTCTTCTCTACTTCGCGCATCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
DB ATCTTCTTCTTCTTCTGCGGTGATCACGCTCATGAC---GTCTGCTGCTGCTGCTGCT 1497
QY 1503 CTCCCGGACACCGCGCGCGGCGGAGCTGGAGGAGATG 1538
DB CTGCGGAGACCAAGGCGGTGCCCATCGAGGAGATG 1533

RESULT 9
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 7.3%; Score 147.2; DB 3; Length 4411529;

834	CGGACCTGTTGGTG-----GCGAGCGAGGAGTCCAAAGCTGGTGCAGCAC	878
897	GAGTCTACTCTGTCCCGACCCCGGCATCGCGGCATCTCTGTGTCGGGATCGGCATC	956
879	CCGTGGCGCAACATCTGCAGCGCAAGTACCGGCCCCAGCTGACCATGGCGATCATGATC	938
957	CACCTTCTTCAGCATGGTTTGGGCATTCACTCCGTGCTCTTCTTACAGCCCTCTCGTGTTC	1016
939	CCCTTCTTCAGCAGCTGACGGGGATCAACGTATCATGTCTTACGGCCCGGTGCTGTTC	998
1017	AAGAGCCCCGGATTAAAGCAACGACAAACACTTCTTGGGCACCACTTGGCCCGTTTCGGTGTTC	1076
999	GAGACGCTGGGTTTCAAGGGCGAGCGTGCCTCATGTCCGCCGTATCATCGGCCCTGGTC	1058
1077	ACCAAGAGGCTTTTCATCTTGTGTGGGACTTTCTTTCATCGAGGGCGTGGCGGGCGGCCG	1136
1059	AACGTGTTCGACGCTCGTGTCCGTTG---TTCACGCTCGACCGCTGGGTTCGCCGAAG	1115
1137	CTGTGTGCTGGCAGCACGGCGGGATAATCTCTCTCCCTCATCGGCCCTCGCGCCGGGCTC	1196
1116	CTGTTCTGCAGGGCGGACGCGATGCTCTGACCGACGCTGTGTGGGACCTTGATC	1175
1197	ACCGTC---GTCCGCGACAGCACCCGACGCGCAAGATACCTTTGGGCCATTCGGCCTAAGCATC	1253
1176	CGCGTCAAGTTCGSGACGAGCGGCGTGGGGGAGATGCCAAGGGGTACGCGGGCGGGTG	1235
1254	GC---CTCCACCTCGCTACGTGCGCTTCTTCTCCATCGSCCTTGCGCCCATCAGTGG	1310
1236	GTGCTCTTCTATCGCTCTATGTGGCCGGTTTCGCTGTGCTGTGGGGGCCCTCTGGGGTGG	1295
1311	GTGTACAGCTCGGAGATCTTCCGCTCCAGTGCAGCGCTGGGCTGCTCGCTCGGCGTC	1370
1296	CTGTCGCCAGCGAGATCTTCCGCTGGAGATCAGGCCGCGGGGACAGATCAACGTG	1355
1371	CGCGCCAACCGGTACACAGGGGGTATCTCCATGACCTTCTCTGCTGTGCTGTCCAGGCC	1430
1356	TCGGTGAACATGCTCTTCACTTCGTATCGCGACGCGTTCCTCACCATGCTCTGCCAC	1415
1431	ATCACCATCGGCGGACGTTCTTCTCTACTCTCGGCATCGCGGCTCGCTCTGGGTGTTTC	1490
1416	ATGAGATT---CGGCTCTTCTACTTCTTCGCGGCTGGGTGTGATCATGACCGTCTTC	1472
1491	TTCTPACACCTTACCTCCCGGAGACCCGCGGCGGACGCTGGAGGAGATG	1538
1473	ATCGCGTCTTCTCTCGCGGAGACCAAGAAAGTGCCTCATCGAGAGATG	1520

RESULT 12

RES001 12
US-09-679-686B-9
; Sequence 9, Application US/09679686B

; PATEIL NO. 6624343
; GENERAL INFORMATION:

```

: APPLICANT: Allen, Stephen M.
: APPLICANT: Lightner, Jonathan E.
: APPLICANT: Rafalski, J. Antoni
: APPLICANT: Thorpe, Catherine J.
: TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
: FILE REFERENCE: BB1160 US NA
: CURRENT APPLICATION NUMBER: US/09/679,686B
: CURRENT FILING DATE: 2003-01-16
: PRIOR APPLICATION NUMBER: 60/081,131
: PRIOR FILING DATE: 1998-04-09
: PRIOR APPLICATION NUMBER: PCT/US99/07561
: PRIOR FILING DATE: 1999-04-07
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 9

```

RESULT 13

RESOLI 13
US-09-103-840A-2
: Sequence 2, Application US/09103840A

; PACIFIC NO. 6294328
: GENERAL INFORMATION:

```

; OTHER INFORMATION: n = a, c, g o r t
;
; FEATURE:
;
; NAME/KEY: unsure
;
; LOCATION: (1031)
;
; OTHER INFORMATION: n = a, c, g o r t
;
; FEATURE:
;
; NAME/KEY: unsure
;
; LOCATION: (1059)
;
; OTHER INFORMATION: n = a, c, g o r t
;
; FEATURE:
;
; NAME/KEY: unsure
;
; LOCATION: (1086)
;
; OTHER INFORMATION: n = a, c, g o r t
;
; FEATURE:
;
; NAME/KEY: unsure
;
; LOCATION: (1126)
;
; OTHER INFORMATION: n = a, c, g o r t
;
; US-09-679-688B-9

```

Query Match 6.7%; Score 136; DB 4; Length 1167;

Query Match 0.78; Score 130; SD 1; Length 1107;
Best Local Similarity 51.5%; Pred. No. 6.8e-18;
Matches 310; Conservative 1; Mismatches 291; Indels 0; Gaps 0;

[illegible]

RESULT 13

RESOLI 13
US-09-103-840A-2
: Sequence 2, Application US/09103840A

; PACIFIC NO. 6294328
: GENERAL INFORMATION:


```
QY 157 CTTTCGGCTGGCCATCTCTCCCTCCATGACCTCCATCCCTCTCGGTACGATATCGGG 216
  |||
Db 80 CGTCTTCGTCGTCTCTCGCCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 139
QY 217 TGATGAGGGGGCTCGGTGATACATCAAGAAGGACTTCAACATCAGTGAAGGAGGTGG 276
  |||
Db 140 TTATTGGGGTGGCTTACCTCTTATTGCGCAATGATGATGATGATGATGATGATGATG 199
QY 277 AGGTTCATCGGGCATACTGACCTCTACTCGCTCATCGCTCCCTTCGCGGGGGGGGA 336
  |||
Db 200 AGTGGGTGGTCAGTCCATGATGTTGGGGGTGCGCTCGCGGGTGGCGAGGGGTGGC 259
QY 337 CGTCGGAAGTGAATCGGGGGGGGTACACCATCGTGTGTTGGCGCGGTGATATTCCTCGCG 386
  |||
Db 260 TCTCTTCAAACTGGGGCGGAAAAAGAGCCGTGATGATCGGGCCATCTCTTCGTGCGCG 319
QY 397 GGGSTTCTCATCGGGTTCGCGCTCAACTACGCAATGCTCATGTTCCGGCGCTTCGTGG 456
  |||
Db 320 GTTCGCTGTTCTTCGCGCGCGGCAAAAGCTGAGATCCCTGCTGGTTCCCGTGTGCTGC 379
QY 457 CCGCATCGGGTGGGTACCGCTCATGATCGCGCGGTGTACACCGCGAGGTGTGCG 516
  |||
Db 380 TCGCGCTGGGTGGGTGGCTCCATATACGGCTCGCTGATCTGTGCGAAATCGCC 439
QY 517 CGGCGTGGCGCGTGGCTTCTGACGTGCTTCCCGAGGTGTTCAATCACTTGGGATCC 576
  |||
Db 440 CAGAAAAAATTCGCGGGCAGTATGATTTCCATGTACCAAGCTGATGATCACCATCGGGATCC 499
QY 577 TGCTCGGTGCTGCTGCAACTATGCTTTCTCCCGCTTTCGCGCTGAACTCGGGTGGCGCA 636
  |||
Db 500 TTGCGGCTATCTCTTGACACCGCTTTCAGTAC-----AGCGGGCATGGGCT 550
QY 637 TCATGCTCGGCATCGGGCGGGCGCTCGGTGCTGCTCGCGCTCATGGTGTGGGATGC 696
  |||
Db 551 GGATGCTCGGGTATCATCATTCGCGCGGTTTGTGCTGATCGGGCTTATCTTCCTGC 610
QY 697 CGGAGTCCGCGCTGGTGTGATGAGGAGCGCTCGCGGCGCAAGGTGGTGTGG 756
  |||
Db 611 CGGACAGCCCGCTGGTGTGCGCGCAAAAGCTGCTTTGCGATGCGGAAACGGGTGCTGC 670
QY 757 AGAAGACCTCCGACACGCGGAGGAGGC 784
  |||
Db 671 TGGCGCTGGCGATACCGAGCGCGAAGC 698
```

RESULT 15

```
US-09-252-991A-3992
; Sequence 3992, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3992
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (870)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-3992
```

Query Match 5.1%; Score 102; DB 4; Length 1344;
Best Local Similarity 48.0%; Pred. No. 3.3e-11;

```
Matches 288; Conservative 1; Mismatches 311; Indels 0; Gaps 0;
QY 169 CCATCTCGCTCCATGACCTCCATCCCTCGGTACGATATCGGGGTGATGAGCGGG 228
  |||
Db 77 CCGTGGCTCCGGGGGGCTCCCTCGTCGACGGCTACGTGTGACATCATCGGGGTGG 136
QY 229 CGTGGCTGATCATCAAGAAAGGACTTCAACATCAGTGAAGGAGGTGAGGTTCATCG 288
  |||
Db 137 CGATGCTGCAGATGACCGGGCATTCGAACATGAGCAGTTTCTGGCAGGCAATGATCGCG 196
QY 289 GCATAGTGAACCTTACTCGCTCATGCGCTCCTTCGCGGGGGGGGAGCTCGGACTGA 348
  |||
Db 197 CTTGCGCGCTGTGCGGCAATCTTCTCGGGGGTTTCTCGCGCGGTGCTGACCGACCGCC 256
QY 349 TCGGCGGGGTGACACCATCGTGTTCGCGCGCGTCAATATCTTCGCGGGGGSGTTCTCA 408
  |||
Db 257 TGGGCGCGCGAGGCTGTTCTTCGTCGCGCCAGCTGTTCGCTCGCTCCCTGGCGC 316
QY 409 TGGGTTCCCGCTCAACTAGGCAATGTTTCGGCGCGTTCGTGGCGCGCATCGGGC 468
  |||
Db 317 AGTACGGGTGACACTCGCGCGCGGCTGTTCTGCTCTGCGCTTCTGATCGCGGTGGCG 376
QY 469 TGGGCTACCGCTCATGATCGCGCGGTGTACACCGCGAGGTGTCCCGCGGTTCGGCGC 528
  |||
Db 377 TGGGATCGAGTACCGGTGCGCAACCGCTTCTGCTGGTGGAGTTCCTGCGAAGATATC 436
QY 529 GTGGCTTCTGACGTGTTTCCCGAGGTGTTCACTCAACTTCGGCATTCCTGCTCGGGTACG 588
  |||
Db 437 GGGACCGCGCTGGCGACCCAGCATCTCTGTTTCGCGCGCGCGCTTCGCTACCC 496
QY 589 TCTCGAACTATGTTTCTCCGCTTCGCGCTGAACCTCGGGTGGCGCATCATGCTCGGCA 648
  |||
Db 497 TGGTGGCGACCTGATTCCTGCGCAGCGGGGGCGAGACCGCTTGGCGCTTGGTACTGGCCA 556
  |||
QY 649 TGGGCGCGCGCGCTCCGCTGCTCGCGCTCATGTTCTCGGCATTCGCGAGTTCGCGCG 708
  |||
Db 557 GCACCGCGCGATCGGTGCGGTCTTCTTCTGCTAGCATCGGCACCCCGAGTTCGCGC 616
QY 709 GGTGGTGTGATGAAGGAGACCGCTCGCGGACGCGCAAGTGTGTGCTGGAGAGACCTCG 768
  |||
Db 617 GCTGGTGTGAGCAAGGGCGCGAGCGGAGCGAGCGGCTCATCGCGAGGTTCATCGCGAGGTCTACG 676
```

Search completed: July 2, 2004, 08:30:48
Job time : 199 secs

This Page Blank (uspto)

Db	121	CCGTCGGCCGAGAAAGAGGGCAACGTCGCGTTTCGCTTCGCGCATCTCTCGCCT	180
Qy	181	CCATGACCTCCATCTCTCTCGGTACGATATCGGGGTGATGAGCGGGCGTTCGCTGTACA	240
Db	181	CCATGACCTCCATCTCTCTCGGTACGATATCGGGGTGATGAGCGGGCGTTCGCTGTACA	240
Qy	241	TCAAGAAGACTTCACATCAGTGA CGGGAAGGTGGAGGTTCTTCATGGGCATACTGAACC	300
Db	241	TCAAGAAGACTTCACATCAGTGA CGGGAAGGTGGAGGTTCTTCATGGGCATACTGAACC	300
Qy	301	TTCTACTCGCTCATCGGTCCTTCGCGCGCGGGGGACGTCGGACTGATCGGCGCGCGT	360
Db	301	TTCTACTCGCTCATCGGTCCTTCGCGCGCGGGGGACGTCGGACTGATCGGCGCGCGT	360
Qy	361	ACACCATCGTGTTCGCGCGCGGTGATACCTTCGCGGGGGGTTCCTCATGGGTTCCCG	420
Db	361	ACACCATCGTGTTCGCGCGCGGTGATACCTTCGCGGGGGGTTCCTCATGGGTTCCCG	420
Qy	421	TCAACTACGCATGCTCATGTTTCGCGCGCTTCGTTGGCGCGCATCGGCGTGGGCTACCGC	480
Db	421	TCAACTACGCATGCTCATGTTTCGCGCGCTTCGTTGGCGCGCATCGGCGTGGGCTACCGC	480
Qy	481	TCATGATCGCGCCGGTGTACACCGCCGAGGTGTTCGCGCGGTTCGCGCGTTCCTGA	540
Db	481	TCATGATCGCGCCGGTGTACACCGCCGAGGTGTTCGCGCGGTTCGCGCGTTCCTGA	540
Qy	541	CGTCTGTTCGCGAGGTGTTTCATCAACTTCGGCATCTCTCGGTGATCGTCTCGAAGTATG	600
Db	541	CGTCTGTTCGCGAGGTGTTTCATCAACTTCGGCATCTCTCGGTGATCGTCTCGAAGTATG	600
Qy	601	CTTTCTCCCGCTTCGCGCTCAACCTCGGGGTGGCGCATCATGCTCGGCATCGGCGCGCGC	660
Db	601	CTTTCTCCCGCTTCGCGCTCAACCTCGGGGTGGCGCATCATGCTCGGCATCGGCGCGCGC	660
Qy	661	CGTCCGTGTCTCGCGCTCATGTTTCGCGCATTCGCGGTTCGCGCGTTCGCTGATCGTCA	720
Db	661	CGTCCGTGTCTCGCGCTCATGTTTCGCGCATTCGCGGTTCGCGCGTTCGCTGATCGTCA	720
Qy	721	TGAAGGGAAGCCCTCGCGAGCCCAAGGTGTGTGGAGAAAGACTCTCGACACGCGCGGAGG	780
Db	721	TGAAGGGAAGCCCTCGCGAGCCCAAGGTGTGTGGAGAAAGACTCTCGACACGCGCGGAGG	780
Qy	781	AGCGCGGAGCGCTTCGCGCATCAAGCGCGCGCGCGCATCTCTTGAAGAGCTCGAGC	840
Db	781	AGCGCGGAGCGCTTCGCGCATCAAGCGCGCGCGCGCATCTCTTGAAGAGCTCGAGC	840
Qy	841	GCACGTGTGACCGTCCCAAGAGAGGAGCGGAACGAGAGAGCGGTGTGGAAGGAGC	900
Db	841	GCACGTGTGTGACCGTCCCAAGAGAGGAGCGGAACGAGAGAGCGGTGTGGAAGGAGC	900
Qy	901	TCATCTGTCTCCCGAACCCCGGCCATCGGCGCATCTCTGTTCGGGATCGGCATCCACT	960
Db	901	TCATCTGTCTCCCGAACCCCGGCCATCGGCGCATCTCTGTTCGGGATCGGCATCCACT	960
Qy	961	TCCTTCAGCATGTTGGGCAATTCATCTCGTTCCTTCACAGCCCTCTCGTGTCAAGA	1020
Db	961	TCCTTCAGCATGTTGGGCAATTCATCTCGTTCCTTCACAGCCCTCTCGTGTCAAGA	1020
Qy	1021	GCCCCGGATTAAAGCAACCAACATTCCTGGGCACCACTTCGCGCGCTTCGCTGTCAACCA	1080
Db	1021	GCCCCGGATTAAAGCAACCAACATTCCTGGGCACCACTTCGCGCGCTTCGCTGTCAACCA	1080
Qy	1081	AGAGGCTTTTCACTTGTGGGACTTTCTTCATCGACGGGTCGCGCGCGCGCGCTGT	1140
Db	1081	AGAGGCTTTTCACTTGTGGGACTTTCTTCATCGACGGGTCGCGCGCGCGCGCTGT	1140
Qy	1141	TGCTGGGACACGCGGGGGATTAATCTCTCCCTCATCTGGCCTTCGCGCGCGCGCTCACCG	1200
Db	1141	TGCTGGGACACGCGGGGGATTAATCTCTCCCTCATCTGGCCTTCGCGCGCGCGCTCACCG	1200
Qy	1201	TCGTCGCGCAGACCCCGACGCCAAGATACCTTGGGCGCATTCGGGCTTAAGCATCGCTTCCA	1260
Db	1201	TCGTCGCGCAGACCCCGACGCCAAGATACCTTGGGCGCATTCGGGCTTAAGCATCGCTTCCA	1260

[illegible]

```

RESULT 2
US-10-051-909-21
/ Sequence 21, Application US/10051909
/ Publication No. US20030199217A1
/ GENERAL INFORMATION:
/ APPLICANT: Allen, Steve
/ APPLICANT: Helentjaris, Tim
/ APPLICANT: Hitz, Bill
/ APPLICANT: Kinney, Tony
/ APPLICANT: Tingey, Scott
/ TITLE OF INVENTION: Plant Sugar Transport Proteins
/ FILE REFERENCE: BAI163 US CIP
/ CURRENT APPLICATION NUMBER: US/10/051,909
/ CURRENT FILING DATE: 2002-01-17
/ PRIOR APPLICATION NUMBER: 60/083,044
/ PRIOR FILING DATE: April 24, 1998
/ NUMBER OF SEQ ID NOS: 38
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 21

```


RESULT 3

US-10-437-963-87990
; Sequence 87990, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 87990
; LENGTH: 1897
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86885C.1
US-10-437-963-87990

Query Match 68.6%; Score 1382.8; DB 17; Length 1897;

Best Local Similarity 84.5%; Pred. No. 0;

Matches 1587; Conservative 1; Mismatches 263; Indels 27; Gaps 2;

QY	135	AGAAAGGGAACGTCGCGTTCGCTTCGCGCTCGCGGATCCCTCGGCTCCATGACCTCCATC	194
DB	16	AGAAAGGGAACGCGAGATGCAATTCACCTCGCGCATCTCGGCTCCATGCGCTCCATC	75
QY	195	CTCCTCGGCTACGATATCGGGTGATGAGCGGGGCTCGCTGTATCATCAAGAAGACTTC	254
DB	76	ATTCTCGGCTATGACATCGGGTGATGAGCGGGGATCGCTGTATCATCAAGAAGACTTC	135
QY	255	AACATCAGTAGCGGAAGGTGGAGTTCTATGGGCATATGAACCTTCTATCGCTCATC	314
DB	136	AGATCACTGACCTGACGTTGGAGATCTGATGGGCACTCTGAACATCTACTCGCTCGTC	195
QY	315	GGCTCTTCGCGCGGGGAGACGCGGCTGATGAGTGGGCTGAGTGGGCTGAGTGGGCTG	374
DB	196	GGCTCTTCGCGCGGGGAGACGCGGCTGATGAGTGGGCTGAGTGGGCTGAGTGGGCTG	255
QY	375	GCGCGGTCATATCTTCGCGGGGSGTTCCTCATGGGGTTTCGCGGTCACATACGCCATG	434
DB	256	GCTGCGGCTTCTTCTTCGCGAGAGCTCTACTCATGGGGTTCTTCGCGTACTAGGCCACT	315
QY	435	CTCATGTTGCGCGGCTTCGTGGCGGATCGCGCTGAGCTAGCGCTCATGATCGCGCG	494
DB	316	CTCATGTTGCGCGGCTTCGTGGCGGATCGCGCTGAGTGGGCTGAGTGGGCTGAGTGGG	375
QY	495	GTGTACACGCGGAGTCTCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG	554
DB	376	GTGTACACGCGGAGATCTCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG	435
QY	555	GTGTTTCATCACTTCGGGATCTCGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG	614
DB	436	GTCTCCATCAACTCGGTATCTCTCGGCTACGCTCTCAACTATGCTTTCGCGCGCTG	495
QY	615	CGGCTGAACCTCGGCTGCGGATCATGCTCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG	674
DB	496	CGGCTCTCGGCTGCGGATCATGCTCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG	555
QY	675	CGGCTCATCTGCTCGGATGCGGAGTTCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG	734
DB	556	CGGCTCATCTGCTCGGATGCGGAGTTCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG	615
QY	735	GCGGACGCGGCTGCTCGGAGAGACTTCGACACGCGGAGGAGCGCGCGGAGCGCGCGGAG	794

DB	616	GCCGACCCCAAGGCTGTGTGGAGAGATCGTGACACGCCGAGAGAGCGCTCGAGCGG	675
QY	795	CTGGCGGACATCAAGCGCGCCCGCATCTCTGAGGAGCTCGAGCGGCGACCTGTGTGACC	854
DB	676	CTCGCTGACATCAAGCGCGCGCTGGCATCCAGATGACCTCGAGCGGCGACGTGTGACC	735
QY	855	GTCCCAAGAGAGGAGCGGAACAGAGACCGGGTGTGGAAGAGCTCATCTGTCTCCCG	914
DB	736	GTGTCCAAAGAGAGCGCGCGAGAGGGGCGAGGTTTGAGGGAGCTGTGTGTCTCCCG	795
QY	915	ACCCGCGCATGCGCGCATCTGTGTCTCGGATCGGCATCCACTTCTTCCAGCATGG	974
DB	796	ACCCGCGCATGCGCGCATCTGTGTCTCGGCGGCTGCGGCTCCACTTCTTCCAGCGG	855
QY	975	TTGGGCACTTCACTTCGCTGTCTTTCACAGCCCTCTCTGTGTTCAGAGCGCGGATTAACG	1034
DB	856	TCGGGCGTCTGACTTCGCTGTCTTACAGCCCGCGGTTCACAGAGCGCGGATCAACC	915
QY	1035	AAGCAAAACACTTCTTGGGACACCTTGGCGGTTGGGCTGTCACCAAGAGCTTTTCATC	1094
DB	916	GCGGACGACGACTCTGTGGGACACCTGCGCGCTGCGGCTTCGCCAAGAGCTTTCATC	975
QY	1095	TTGTTGGCGACTTCTTTCATCGACGCGCTCGGCGCGCGCGCTGTGTGTGGGACGACG	1154
DB	976	CTGGTCCGCGGTTCTGTCTCGACCGCGCGCGCGCGCTGTGTGTGACGACGACG	1035
QY	1155	GGGGGATTAATCTCTCTCTCATCGGCTCGGCGCGCGGCTCACCGTGTGGGCGGACG	1214
DB	1036	GGGGGATGTCTTCTCTCTCTCGGCTCGGCGCGCGGCTCACCGTGTGGGCGGCGG	1095
QY	1215	CCGACGCGCAAGATACCTTGGGCGCATCGGCTTAAGCATCGCTCCACCTCGCCTAC	1271
DB	1096	CGGACGCGCGAGTCCGAGCTGGGCGCTGCGGCTGTGGTGTGGTGTTCATCTTGGCTAC	1155
QY	1272	GTGCGCTTCTTCTTCATCGGCTTGGCGGCTTTCAGCGGCTGTAAGTGTGGAGATCTTC	1331
DB	1156	GTGCGCTTCTTCTTCGCTCGGCTCGGCGGCTTTCAGCGGCTGTAAGTGTGGAGATCTTC	1215
QY	1332	CGGCTCAGTGGCGGCTGGCTGTCTGCTGCGGCTCGGCGGCGGCGGCTGCGGCTGCGG	1391
DB	1216	CGGCTCAGTGGCGGCTGGCTGTCTGCTGCGGCTCGGCGGCGGCGGCTGCGGCTGCGG	1275
QY	1392	GGCGTCTATCTCCATGACCTTCTGCTGTCTGCTGCTCAAGAGGCTATCAGCTCGGCGG	1451
DB	1276	GGGCTCTCTCCATGACCTTCTGCTGCTTCCAGCGGCTATCAGCTCGGCGGCGGCTTC	1335
QY	1452	TTCTCTCTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1511
DB	1336	TTCTCTCTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1395
QY	1512	ACCGCGCGCGGACGCTGGAGAGATGAGCAAGCTGTTGCGGCGACACGCGCGCGGCTCG	1571
DB	1396	ACCGCGCGCGGACGCTGGAGAGATGAGCAAGCTGTTGCGGCGACACGCGCGGCTCG	1455
QY	1572	GAATCAGACGCGGCGGAGAGAA-----GAAGGTG	1607
DB	1456	GAAGCAGAGGACTTCGCGCGCTACAGGGAAGACTCTTGGCAACTTCTTCATGAAGGTG	1515
QY	1608	GAAATGCGCGGCTAACTGATCAAACTAAACCGCAAAATCAACAAATCTTAAGGGTTTC	1667
DB	1516	GAAATGCGCGGCTAACTGATCAAACTAAACCGCAAAATCAACAAATCTTAAGGGTTTC	1575
QY	1668	TTGCAAAAAGCTGT	1727
DB	1576	TTGCAAAAAGCTGT	1635
QY	1728	CGCTGATCGGCTGT	1787
DB	1636	CGCTGATCGGCTGT	1695
QY	1788	AGACTTCTTAAATCATCTTCAAGTACATGATTTTATTTTGTCTTTTGTCTTTTGTCT	1847
DB	1696	AGACTTCTTAAATCATCTTCAAGTACATGATTTTATTTTGTCTTTTGTCTTTTGTCT	1755


```

Db      1625  CAAAGTTGGCAGATCCGATGCGAAGACTTGCGCTGTTATCGTCTCGGCTAGCTAGCTAGCTGCC 1684
QY      1701  A-----AGTAGTAGCAGCAACAGTGGGAAGATTGCTGATCGGCG-----GTTGCT 1745
Db      1685  ACAAGGCCACATAGATGACGAAGTAGCGTGGGAAGATTGCTGATCGGCGCGAGCTGCC 1744
QY      1746  GGAGGAGCGACGGCCGCG 1762
Db      1745  GGAGGCGGACGGCAAGC 1761

RESULT 6
US-10-425-114-30198
; Sequence 30198, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30198
; LENGTH: 1093
; TYPE: DNA
; ORGANISM: Oryza sativa cypress
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-OSROCCYP002H08_FLI
US-10-425-114-30198

Query Match          51.6%; Score 1039.8; DB 13; Length 1093;
Best Local Similarity 96.9%; Pred. No. 6.9e-254;
Matches 1059; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

QY      35  CTCGACCGCCACTACTGTACAGCGGCCAGACGAGCTCTCTCTCTCTGACACCGGA 94
Db      1  CTCGACCGCCACTACTGTACAGCGGCCAGACGAGCTCTCTCTCTGACACCGGA 60
QY      95  GATGCGTTCGCGCGCTCTCGCGAGCGCTCGCGCGAAGAGGCAACCTCGGTT 154
Db      61  GATGCGTTCGCGCGCTCTCGCGAGCGCTCGCGCGAAGAGGCAACCTCGGTT 120
QY      155  CGCTTCGCTCGGCATCTCGCTCCATGACCTCCATCTCTCTCTGCTACGATATCG 214
Db      121  CGCTTCGCTCGGCATCTCTCGCTCCATGACCTCCATCTCTCTCTGCTACGATATCG 180
QY      215  GGTGATGAGCGGGGGTCTGCTGATACAAAGAGGACTTCAACATCAGTGACGGGAAGT 274
Db      181  GGTGATGAGCGGGGGTCTGCTGATACAAAGAGGACTTCAACATCAGTGACGGGAAGT 240
QY      275  GGAGGTTCTCATGGGCATAGTAACTCTACTCTGCTCATCGGCTCTCTCGCGCGGGCG 334
Db      241  GGAGGTTCTCATGGGCATAGTAACTCTACTCTGCTCATCGGCTCTCTCGCGCGGGCG 300
QY      335  GAGCTCGGACTGAGTACGGCGCGGTACACCATCGTGTTCGCCGCCCTCATATCTTCGC 394
Db      301  GAGCTCGGACTGAGTACGGCGCGGTACACCATCGTGTTCGCCGCCCTCATATCTTCGC 360
QY      395  GGGGGSGTTCTCATGGGGTTCGGCGTCAACTAGCGCATGCTCATGTTTCGGCGCGCTCGT 454
Db      361  GGGGGSGTTCTCATGGGGTTCGGCGTCAACTAGCGCATGCTCATGTTTCGGCGCGCTCGT 420
QY      455  GGCGGCGATCGCGGGGTACGCGCTCATGATCGCGCGGTGTACACCGCCGAGGTGTC 514
Db      421  GGCGGCGATCGCGGGTACGCGCTCATGATCGCGCGGTGTACACCGCCGAGGTGTC 480
QY      515  GCCGCGTTCGGCGCGGTGCTTCTGACGTCGTTCCCGGAGGTGTTTCATCAACATCGGCA 574

```

Db 55 GATGGCTTCGACGAGCTCGAAAGGCGCTCGAGCCAGGAGAGGCGCAACGTCAGTA 114
Qy 155 CGCCTTCGCTCGGCATCCTCGCTCCATGACCTCCATCTCTCGCTGCGATATCGG 214
Db 115 TGCCCTCATATGTGCACTTCGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 174
Qy 215 GGTATGAGCGGGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 274
Db 175 GGTATGAGTGTGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 234
Qy 275 GGAGTTCCTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 334
Db 235 GGAGATCTGATCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 294
Qy 335 GACGTCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 394
Db 295 GACGTCGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 354
Qy 395 GGGGSGTTCCTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 454
Db 355 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
Qy 455 GGGCGGATCGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 514
Db 415 GGGCGGATCGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474
Qy 515 GCGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 574
Db 475 GCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 534
Qy 575 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634
Db 535 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
Qy 635 CATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 694
Db 595 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
Qy 695 GCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 754
Db 755 GCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
Qy 755 GGAGAGAGCTTCCGACACGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 814
Db 715 AGAGAGAGCTTCCGACACGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 774
Qy 815 CGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 871
Db 775 GCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834
Qy 875 CGGAG 931
Db 835 CGGCGGTGAGTGTGAGGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAG 894
Qy 932 CATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991
Db 895 CATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
Qy 992 CGTCTTCTACAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051
Db 955 CGTCAAGTACAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
Qy 1052 GGGACACACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1111
Db 1015 GGGCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074
Qy 1112 CATGACAGCGGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1171
Db 1075 GCTGACCGCGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
Qy 1172 CTTATCGGCT 1231
Db 1135 GCTCATCTGCT 1194

Qy 1232 TTGGGCTTCGCGCTTAAGCATCGCTTCAACCTTCGCTTACGCTGCTTCTTCTTCCATCGG 1291
Db 1195 GTGGGCGCTGCGCTTGTGATCGCTCAACCTTGTCTACATCGCTTCTTCTTCCATCGG 1254
Qy 1292 CTTGGGCGGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1351
Db 1255 CTTGGGCGGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
Qy 1352 GGGCT 1411
Db 1315 GGGCT 1374
Qy 1412 CTTGCT 1471
Db 1375 CTTGCT 1434
Qy 1472 CGGCT 1531
Db 1435 CGGCT 1494
Qy 1532 GGGATGAGCAAGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1585
Db 1495 GGAGATGGCAAGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1554
Qy 1586 AGCCAGGAG 1626
Db 1555 CGCAGCAAGGAG 1595

RESULT 8

US-10-051-909-19
; Sequence 19, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-10-051-909-19

Query Match 49.8%; Score 1005; DB 14; Length 1914;
Best Local Similarity 79.2%; Pred. No. 6,6e-245;
Matches 1220; Conservative 1; Mismatches 311; Indels 9; Gaps 2;
Qy 95 GATGGCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 154
Db 55 GATGGCTTCGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 114
Qy 155 CGCCTTCGCT 214
Db 115 TGCCCTCATATGTGCACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 174
Qy 215 GGTATGAGCGGGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 274
Db 175 GGTATGAGTGTGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 234
Qy 275 GGAGGTTCTCATGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 334
Db 235 GGAGATCTGCT 294

Db 548 GGTCGGCCCGCGCTTCTTCTACAGTCTTCCACGAGTGTTCATCAATGTGGCATCTC 607
Qy 579 CTCGGGTACGTCGAACTATGCTTTCTCCGCTTGGCGCTGAACTCTGGGTGGCGATC 638
Db 608 CTGTGCTACGTCGAACTATGCTTTCTCCGCTTGGCGCTGAACTCTGGGTGGCGATC 667
Qy 639 ATGCTCGGATCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
Db 668 ATGCTCGGATCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727
Qy 699 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
Db 728 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
Qy 759 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
Db 788 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 847
Qy 819 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 878
Db 848 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907
Qy 879 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 938
Db 908 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 967
Qy 939 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
Db 968 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027
Qy 999 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058
Db 1028 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087
Qy 1059 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1118
Db 1088 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147
Qy 1119 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1178
Db 1148 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1207
Qy 1179 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1238
Db 1208 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267
Qy 1239 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
Db 1268 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1327
Qy 1299 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1358
Db 1328 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1387
Qy 1359 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1418
Db 1388 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1447
Qy 1419 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1478
Db 1448 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1507
Qy 1479 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1538
Db 1508 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1567
Qy 1539 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1598
Db 1568 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1627
Qy 1599 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632

Db 1628 AAGCGAAAGTAGGGGAGATGAACATAGTAGCTA 1661
RESULT 10
US-10-051-909-27
; Sequence 27, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B11163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Triticum aestivum
; US-10-051-909-27
Query Match 48.7%; Score 981.6; DB 14; Length 1872;
Best Local Similarity 77.4%; Pred. No. 5.8e-239;
Matches 1188; Conservative 1; Mismatches 345; Indels 0; Gaps 0;
Qy 99 GCTTCGCGCGCTGCGGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 158
Db 128 GCTCGCGCTCCCGAGCG 187
Qy 159 TTGCGCTGCGCGCGCTGCG 218
Db 188 TTGCGCTGCGCGCGCTGCG 247
Qy 219 ATGAGCGCGCGCTGCG 278
Db 248 ATGAGCGCGCGCTGCG 307
Qy 279 GTTCTCATCG 338
Db 308 ATCATGATGCG 367
Qy 339 TCGGACTGGATCG 398
Db 368 TCGGACTGGATCG 427
Qy 399 GSGTTCTCATGGGTTGCG 458
Db 428 TCCTTCTCATGGGTTGCG 487
Qy 459 GGCATCG 518
Db 488 GGAATCG 547
Qy 519 GCGTCG 578
Db 548 GCGTCG 607
Qy 579 CTCGGGTACGTCGAACTATGCTTTCTCCGCTTGGCGCTGAACTCTGGGTGGCGCATC 638
Db 608 CTGCGGTACGTCGAACTATGCTTTCTCCGCTTGGCGCTGAACTCTGGGTGGCGCATC 667
Qy 639 ATGCTCGGATCGGCG 698
Db 668 ATGCTCGGATCGGCG 727
Qy 699 GAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758

Db 728 GAGTCTCTCGTGGTCTCATGAAGGCGCGCTCGGAGCGCCAGCGGCGTCTGGCC 787
Qy 759 AAGACCTCCGACACGCGGAGAGCGCGGAGCGCTGCGGACATCAAGGCGCGCC 818
Db 788 AAGACCTCCGACACGCGGAGAGCGCGTGGAGCGCTTGACACAGATCAAGGCTGCGGCC 847
Qy 819 GGCATCCCTAGGAGCTCGAGCGGAGCGTGGTACCGTCCCAAGAGAGGAGCGGAAC 878
Db 848 GGCATCCCTAGGAGCTCGAGCGGAGCGTGGTACCGTCCCAAGAGAGGAGCGGCGAG 907
Qy 879 GAGAAGCGGCTGGAAGAGCTCATCTGTCCCGGACCGCGGCGATCGGCGATCCTG 938
Db 908 GAGAAGCAGGTGTGGAAGAGCTCATCTTTTCGCGAGCGCGGCGATCGGCGCATCTG 967
Qy 939 CTGTCCGGATCGGATCCACTTCTTCAGCATCGGTTGGCATCTACTCGGTCGCTTC 998
Db 968 CTCGCGGCTCGGATCCACTTCTTCAGCAGCGAGCGGCTCCGACTCCGTCGCTC 1027
Qy 999 TACAGCCCTCGTGTTCAGAGCGCGGATTACAGAACAGCAACACTTCTTGGGCAOC 1058
Db 1028 TATAGCCACCGGTGTTCAGAGCGGCGCATCACCGGACACACCTGCTCGGCGCC 1087
Qy 1059 ACTTGGCGGTTCGGTGTACCAAGAGCTTTTCATCTTGTGTGGGCACTTCTTCATCGAC 1118
Db 1088 ACATGCGCATGCGGCTCATAGACGCTCTTCATCTGTGTGGCAGGTTCCAGCTCGAC 1147
Qy 1119 GGCCTCGGCGCGCGCTGTCTGGGAGCAGCGGCGGATAATCTCTCCCTCATC 1178
Db 1148 CGGTTCGGAGCGCGCGCTCTGTACACAGCAGCGGCGCATCTGCTGCTGTCTATC 1207
Qy 1179 GGCCTCGGCGCGCGCTCACGCTCGTGGCGAGCACCCGACGCAAGATACCTTGGGCC 1238
Db 1208 GGCCTCGGAGCGGCTCACGCTCGTGGTGGCACCGGAGCGCAAGGTCCTGGGCC 1267
Qy 1239 ATCGGCTAAGCATCGCTCCACCTCGCTACGTCGCTCTTCTTCATCGGCTTGGC 1298
Db 1268 ATCGGCTGTGCATCGTGTCCATCTTGGCTACGTCCTTCTTCTCCATCGGCTCGGG 1327
Qy 1299 CCCATCAGCTGGGTGTACAGCTCGGAGATCTTCCGCTCCAGGTGCGCGCGCTGGGTGC 1358
Db 1328 CCCCTCACCACGCTGTACACCTCGGAGTCTTCCACTGCGGTGCGCGCGCTGGGTTC 1387
Qy 1359 TCGCTCGGCTGCGCGCAACCGGCTCACAGCGGCTCATCTCATGACCTTCTCTGTG 1418
Db 1388 CGCTTGGGCACTCATGAACCGGCTCACAGCGCGGCTCTCCATGCTCTTCTCTGTC 1447
Qy 1419 CTGTCCAGGCGCATCACCATCGGCGAGCTTCTCTCTACTCCGCGATCGCGGCTC 1478
Db 1448 TTGTCCAGGCGCATCACCATCGGCGAGCTTCTCTCTACTCCGCGATCGCGGCGATA 1507
Qy 1479 GCCTGGGTGTTCTTCTACACCTACTCTCCGAGACCGCGCGGACGCTGGAGAGATG 1538
Db 1508 GGATGGAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1567
Qy 1539 AGCAAGCTTTTCGGGACACCGCGCTCGGAATCAGACGAGCGGCGGAGAGAG 1598
Db 1568 GGGAGCTTTTCGGGATGAGGACGCGCTGCAAGCGGCGGAGAGCGGCGGAGAGAGC 1627
Qy 1599 AAGAGGTGGAATGCGCGCACTAACTGATCAA 1632
Db 1628 AAGCGAAGTAGGGAGATGAAGTAGCTA 1661

RESULT 11
US-10-425-114-28112
; Sequence 28112, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 28112
; LENGTH: 1704
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4746-010-A4_FLI
US-10-425-114-28112

Query Match 38.7%; Score 779.6; DB 13; Length 1704;

Best Local Similarity 71.4%; Pred. No. 1.3e-187;

Matches 1044; Conservative 1; Mismatches 405; Indels 12; Gaps 1;

Qy 152 GTTCGGCTTCGCTCGCCATCTCGCTCCATGACCTCCATCCTCCTCGGTACGATAT 211
Db 136 GTACGGCTTCGCTCGCCGCTCTCTGCTTCATGAACTCCATCCTCTTGGTACGATCT 195
Qy 212 CGGGGTGATGAGCGGGCGTCTGTATCATCAAGAGAGACTTCAACATCAGTGACGGAA 271
Db 196 GTCCGTGATGAGCGGAGCGAGTATTATGAAGAGAGACCTCAAGATCAGGACACGCA 255
Qy 272 GTTGGAGTTCATGCGGATACGACCTCTACTCGCTCATCGGCTCCTTCGCGCGG 331
Db 256 GATCGAGATCTTCGCGGCGATCATCAATCTACTGCTCTCGGCTCGCTCGCGCGG 315
Qy 332 GCGGAGCTCGGATGATCGCGCGGCGGTACACCATCTGTGTTCGCGCGCGTTCATATCTT 391
Db 316 CGGAGCTCCGACTGCTCGCGCGGCGGTACACCATGCTGTGTTCGCGCGCGCATCTTCT 375
Qy 392 CGCGGGGSGTCTCTCATGGGTTCGCTCACTAGCCATGCTCATGTTGGCGGCTT 451
Db 376 CGCGGGCGGCTCATCATGGGCTCGCGCGGAGTACGCGATCCTCATGCTTGGCGGCTT 435
Qy 452 CGTGGCGCGCATCGGGTGGGTACGCGCTCATGATCGCGCGGTGTACACCGCGAGGT 511
Db 436 CGTGGCGGCGTTCGGGTGCGATACGCGCTCATGCGACCCGCTGACACCGCGAGGT 495
Qy 512 GTTCGCGGCTCGCGCGTGGTTCGTGAGTCTGTTCCGAGAGGTTCATCAACTTCGG 571
Db 496 CTCGCCAGCTCGCGCGGCGTGTCTACGCTCTTCCGAGAGGTTCATCAACACACGG 555
Qy 572 CATCTGCTCGGTAAGTCTCGAACTATGCTTTCCTCGCTTCGCGCTGAACTCGGCTG 631
Db 556 GGTGCTCTCGGGATGCTCTCAACTACGCTTCACGCGCTCCCGCTCGACCTCAGCTG 615
Qy 632 GCGCATCATGCTCGGCGATCGCGCGCGGCTCGCTCGTGTCTCGGCTCATGCTGCTCG 691
Db 616 GCGGCTCATGTTCTCTCGTGGCGGCTCGCGCTATCTTCTCGCACTGGGGGTCTCTCG 675
Qy 692 CATCGCGAGTTCGCGCGGTGCTGTATGAGGAGCGCTTCGCGGAGCGCAAGGTGT 751
Db 676 CATCGCGAGTTCGCGCGGTGCTGTATGAGGAGCGCTTCGCGGAGCGCAAGGTGT 735
Qy 752 GTTGAGAGAGACTTCGACACCGCGGAGGAGCGCGGAGCGCTTGGCGGATCAAGGC 811
Db 736 GCTGGCCAGAGACTCCAACTCCCGCGGAGCGCGGAGCGGCTCGCGGATCAAGAA 795
Qy 812 CGCGCGCGCATCTCTGAGG-----AGTTCGAGCGGAGCGGTGAGCGTCC 859
Db 796 TGCATTCGGCTCCAGAGCGGCTCGCGCAACACGACGAGTGTGTGTGTGTGTGTGTGT 855
Qy 860 CAAGAGAGGAGCGGAAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 919
Db 856 CAGGAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 915
Qy 920 GGCATGCGGCGCATCTCTGCTCGGGGATCGGCGATCCACTTCTTCCAGCATGCGTGGG 979

Db 916 GCCCGTCGCCGCACTACTACGCTCGCTCGGCTCCAGTTCTTCAGAGGCTCCGG 975
Qy 980 CATTCACCTCGGTCGCTTCTACAGCCCTCTCGTGTTCAGAGCCCGGATTAAAGAA 1039
Db 976 CATCGACTCGGTCGCTGTACAGCCCGCGGTGTTCAGAGGCTCGGCTTAGGTGAA 1035
Qy 1040 CHAAACACTTCTTGCGGCACTTGGCGCTTGGTGTACAGAGGCTTTTCACTTGT 1099
Db 1036 CAACAACCTTTTGGCGGCCACCATGGCGGTGGCGCGACCAAGAGCTGTTCATCTGGT 1095
Qy 1100 GCGACACTTCTTCACTGACGCGCTCGGCGCGCGCGCTGTGTGGCGACAGCGGCGG 1159
Db 1096 GGCACAGTCTTCTCGACGCGCTCGGCGGAGAGCGCTGTCTCAACAAGCGGCGG 1155
Qy 1160 GATAATCTCTCTCCATCGGCTCGGCGCGCGGCTCACCGTGTGGCGCAGACCCCGA 1219
Db 1156 GATGCTGTGTCTGTGTGACGCTGGCTCGGCTCGGCGCTTGGCGGCCATCGACCGGCTCCGAGA 1215
Qy 1220 CGCCAAAGATACCTTGGGCCATCGGCTTAAGCATGCTCTCCACCCCTCGCTACGTCGCTT 1279
Db 1216 GGCACAGGAGCTCGCTGGCGGCGCTGAGCATCGCGCGGTTCTCAGCTTCTGTGGCTC 1275
Qy 1280 TTCTCCATCGGCTTGGCGGCCATCACGCTGGGTGACGCTCGAGATCTTCCGCTCCA 1339
Db 1276 TTCTCCATCGGCTTGGCGGCCATCGCTGGGTGACAGCTCGAGATATTCACACTGCG 1335
Qy 1340 GGTGCGCGCTGGCTGCTCGCTCGGCTCGCGCCCAACCGCTCACAGCGGCTCAT 1399
Db 1336 GCTCGCGCGGAGGCTCGGCACTCGGACGCGGATGAACCGGTCATGAGCGGACCAT 1395
Qy 1400 TTCTCATGACTTCTCTGCTGCTTCAAGGCCATCACATCGCGCGGCTTCTTCTCTA 1459
Db 1396 CACCATGCTCTCATCTGCTCTACAAGGCCATCACCTTTCGCGGAGCTTCTACTTGA 1455
Qy 1460 CTCGCGCATCGCGCTCGCTCGCTGGTGTCTTCTACACTTCTTCCGAGACCCCGG 1519
Db 1456 TCGAGGCACTTGTGCTCGCGGTGCTGTCTATGTTCTTCTCTCGCGGAGACGAGGG 1515
Qy 1520 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGCGCACACGCGCGCTCCGGAATCAGA 1579
Db 1516 GAGCAACCTAGAGACACCGAGAGCTCTTCGGCGGTGGCGACCAAGAGGACAGGA 1575
Qy 1580 CGAGCCAGCAAGGAGAGAAG 1601
Db 1576 AGATGGACACGATGCGCAGAAG 1597

RESULT 12

US-10-437-963-78515
; Sequence 78515, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 78515
; LENGTH: 1809
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MFT4530_78314C.1
US-10-437-963-78515

Query Match 37.6%; Score 759.4; DB 17; Length 1809;
Best Local Similarity 73.0%; Pred. No. 1.9e-182;
Matches 992; Conservative 1; Mismatches 357; Indels 9; Gaps 1;
Qy 207 GATATCGGGGTATGAGCGGGCTCGCTGTATCATCAAGAAGGACTTCAACATCATGAC 266
Db 220 GACATCTCGGTATGAGCGCGCGCAGAAAGTTTCATGAAGAGACCTGAACATCTCGAC 279
Qy 267 GGAAGGTGAGGTTTCTATGGGCTACTGAACCTTACTCTCTATCGCTCATCGGCTCTTCGG 326
Db 280 GCGAAGGTGAGGTTCTCGCGGCTCATCAACATCTACTCTCTCTCGGCTCTCGCTCGG 339
Qy 327 GCGGGCGGACGTCGACTGGATCGCGCGGCTACCATCGTGTTCGCGCGGCTCAT 386
Db 340 GCGGGCGGACGTCGACTGGATCGCGCGGCTACCATCGTGTTCGCGCGGCTCAT 399
Qy 387 TTCTTCGCGGGGTTCTCTATGGGTTTCCTGAGCTTCTTCCGAGGTGTTCATCAAC 446
Db 400 TTCTTCGCGGGGCTCATATGGGCTCGCGCGGCTTGTCTCACGTCTTCCCGAGGTGTTCATCAAC 459
Qy 447 CGCTTCGTCGCGGCTACGCTGAGCTACGCTCATGATCGCGCGGCTACACCGCC 506
Db 460 CGCTTCGTCGCGGCTACGCTGAGCTACGCTCATGATCGCGCGGCTACACCGCC 519
Qy 507 GAGTGTCTCGCGGCTCGGCTTCTGAGCTTCTTCCGAGGTGTTCATCAAC 566
Db 520 GAGTGTCTCGCGGCTCGGCTTCTGAGCTTCTTCCGAGGTGTTCATCAAC 579
Qy 567 TTGCGCATCTGCTCGGCTACGCTGAGCTTCTTCCGAGGTGTTCATCAAC 626
Db 580 GCGCGGCTCTCTCGGCTACGCTTCAACCTTCTTCCGAGGTGTTCATCAAC 639
Qy 627 GGTGCGCATATGCTCGGCTTCTGAGCTTCTTCCGAGGTGTTCATCAAC 686
Db 640 GGTGCGCATATGCTCGGCTTCTGAGCTTCTTCCGAGGTGTTCATCAAC 699
Qy 687 CTGCGCATCTGCTCGGCTACGCTGAGCTTCTTCCGAGGTGTTCATCAAC 746
Db 700 CTGCGCATCTGCTCGGCTACGCTTCAACCTTCTTCCGAGGTGTTCATCAAC 759
Qy 747 GTGCTCTCGGAGACCTTCCGAGCTTCTTCCGAGGTGTTCATCAAC 806
Db 760 GTGCTCTCGGAGACCTTCCGAGCTTCTTCCGAGGTGTTCATCAAC 819
Qy 807 AAGCGCGCGCTCGGCTTCTGAGGAGCTTCTTCCGAGGTGTTCATCAAC 857
Db 820 AAGCGCGCTCGGCTTCTGAGGAGCTTCTTCCGAGGTGTTCATCAAC 879
Qy 858 CCAAGAGAGGAGCGGAAACAGAGCGGCTTTCGAGAGGCTCATCTGTCGCGGAC 917
Db 880 CACAAGAGAGGAGGCTTCCGAGGCTTCTGAGGAGGCTTCTGTCGCGCGGAG 939
Qy 918 CCGGCGCATCGGCGCATCTGCTTCCGAGGCTTCTTCCGAGGTGTTCATCAAC 977
Db 940 CCGGCGCTGCGCGCATCTCATCGCTTCTTCCGAGGTGTTCATCAAC 999
Qy 978 GGAATTCATCTCGTCTTCTTACAGCTTCTGAGGCTTCTTCCGAGGTGTTCATCAAC 1037
Db 1000 GGAATTCATCTCGTCTTCTTACAGCTTCTGAGGCTTCTTCCGAGGTGTTCATCAAC 1059
Qy 1038 GAGAAACACTTCTTGGGCACTTCTGCGGCTTCTTCCGAGGTGTTCATCAAC 1097
Db 1060 GATCTCGAGCTTCTTACAGGCTTCTTCCGAGGTGTTCATCAAC 1119
Qy 1098 TTGCGGACTTCTTCTTACAGGCTTCTTCCGAGGTGTTCATCAAC 1157
Db 1120 GTGGCAAGCTTCTTCTTACAGGCTTCTTCCGAGGTGTTCATCAAC 1179
Qy 1158 GGAATTCATCTCTCGTCTTCTTACAGGCTTCTTCCGAGGTGTTCATCAAC 1217
Db 1180 GGAATTCATCTCTCGTCTTCTTACAGGCTTCTTCCGAGGTGTTCATCAAC 1239

1218 GAGCCAAAGATACCTTTGGGCCATCGCCCTAAGCATCGCTCCACCCTCGCTACGTCGCC 1277
1240 GAGGCCAGCAGCAGCGCTGTGGTCTGAGCATCGCGATGCTGTGTCTGTGGCG 1299
1278 TTCTTCTCATCGCCCTTGGCCCATCAAGTGGGTGTACAGTCTCGAGATCTTCCGCTC 1337
1300 TCCTTCTCATCGGATGGCCGATCGCGTGGGTGTACAGTCTCGAGATCTTCCGCTG 1359
1338 CAGTGGCGCGTGGGCTCTCGCTCGGCGTTCGCCCAACCGCGTCAACGAGCGCGTC 1397
1360 CGCTGGCGCGAGGATCGCGCTCGGACCGCGATGAACCGGTCTGTAGCGCGGC 1419
1398 ATCTCATGACCTTCTGTCTGTCGATGTCGAGGATCAACATCGCGCGAGCTTCTTCTC 1457
1420 GTGAGCATGCTCTCATCTGCTCTACAGGCAATCACTTCTCGCGGAGCTTCTACCTC 1479
1458 TACTCGGCATCGCGCGCTCGCTGGGTGTTCTTCTACACCTACCTCGCGAGACCGC 1517
1480 TAGCGGCAATAGTCGCGCGGTGGGTGTTCAAGTCTTCTTCTGCGGAGACGCA 1539
1518 GGCGGACGCTGGAGGAGATGAGCAAGCTTTCGGCGAC 1556
1540 GGCAGGAGCTGGAGGACACCGTCAAACTCTTCGGCGGC 1578

RESULT 13
US-10-425-114-30135
; Sequence 30135, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30135
; LENGTH: 2203
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-OSFLCYP011G10_FLI
US-10-425-114-30135

Query Match 36.7%; Score 739.6; DB 13; Length 2203;
Best Local Similarity 69.1%; Pred. No. 2.3e-177; Mismatches 450; Indels 9; Gaps 1;
Matches 1028; Conservative 1

104 CGCGCGGTGCGGAGCGCTCGCGCGGAGAGAGAGGCAACGTCGCGTTCGCTTCGC 163
218 CGCGCTCCGTGGCGGTGGCGGAGAGCGCGCCCATCAACAGTACGCTTCGG 277
164 GTGGCCCATCTCGCTCCATGACCTCATCTCTCGGTACGATATCGGGTGTAG 223
278 CTGGCTCTCTCGCTCCATGAACTCGTCTCTCGGTATGATATCTCGGTATGAG 337
224 CGGGCGTGTGTACATCAAGAGGACTTCAACATCAGTACGAGGAGGTGAGGTTCT 283
338 CGGCGGAGATATTCATGAGAGGAGTTTGAAGATCACGACACGAGATCGAGATCCT 397
284 CATGGGCATGTAACCTTACTCGCTACATCGGTCTCTCGCGCGGCGGAGCGTCCGA 343
398 CGCGCGGTATCAACATCTACTCGTGTTCGCTCTCGGTCTCGTGGCGGCGCATGACGCCGA 457
344 CTGATCGCGCGGTGACACATCGTGTTCGCGCGCTCATATCTTCGCGGGSGT 403
458 CTGGCTGGCGCGGTACACATGGTGTCTCGCGCGCCCATCTTCTTACGGCGCGCT 517

404 CTTATGGGGTTCGCGCTCAACTACGCGCATCTCATGTTTCGCGCGCTTCGTTGGCGGCGAT 463
518 CTTATGGGGTTCGCGCGCAACTACGCGCTTCCTCATGCGCGCGCTTCGTTGGCGGCGAT 577
464 CGCGTGGGCTACGGCTCATGATCGCGCGGTGTACACCGCGAGGTGTGCGCGCGCTC 523
578 CGCGTGGGCTACGGCTCATGATCGCGCGGTGTACACCGCGAGGTGTGCGCGCGCTC 637
524 CGCGTGGGCTTCGTGAGCGTGTTCGCGGAGGTGTTCATCAACTTCGCGCATCTCTCTCGG 583
638 GCGCGCGGCTTCCTCACTCTGTTCCCGAGGTGTTCACAAACAGCGGCACTCTCTCTCGG 697
584 GTACGCTCGAACTATGCTTTCGCGCTTCGCGTGAACCTCGGTGGCGCATCATGCT 643
698 CTACGCTCTCAACTTCGCGCTTCGCGCGCTTCGCGTCCACCTGAGCTGGCGCGCATGTT 757
644 CGGCATCGCGCGCGCGCTGCTGCTCGCGCTCATGTTGCTCGCATGCGCGAGTC 703
758 CTTGTCGCGCGCTGCGCGCATCTTCTCGCATCGCGCTCTCTCGCATGCGCGAGTC 817
704 GCGCGGTGGTGTGATGAAAGGAGCGCTCGCGAGCGCAAGGTGTGCTGGAGAAC 763
818 GCGCGGTGGTGTGATGAAAGGAGCGCTCGCGAGCGCGCGCTCTCTCTCAAGAC 877
764 CTCGACACGCGGAGGAGCGCGGCGCTGCGCGCATCAAGGCGCGCGCGGAT 823
878 CTCGACTCGCGGAGGAGCGCGGCTCTCTGACATCAAGAAAGCGCTCGGAT 937
824 CCTGAG-----GAGCTCGACGCGGAGTGTGTCGCTGCTGCGCGAGAGGAGCGG 874
938 CCGCGGAGCGGCTCGGATGCGGAGACGTGTGCTGCTGCGCGCGGCAACAGGCTC 997
875 AAACGAGAGCGGCTGTGAGGAGCTCATCTGTCGCGGAGCGCGCGGATCGCGGAT 934
998 ACAGGAGAGGCGGTGTGAGGAGCTGTGCTCAACCGGAGCGCGCGCTCGCGCGAT 1057
935 CTTGTCGCGGATCGGATCGGATCTTCTCAGCATGCTTGGGCACTTCACTCGCTGT 994
1058 GTCTGTCGCGGCTCGCGCTGATGTTTATCAGCAAGCGCACCGCGCTCGCTCGT 1117
995 CTTTACAGCCTCTCGTTCAGAGCGCGGATTAAGCAAGCAACAACTTCTTGGG 1054
1118 GATGTACAGCGCGGTGTTCAGCGCGCGGCTTAAGTCAAGAGCAACTCTCGG 1177
1055 CACCACTTGGCGTTCGCTGTCACCAAGAGCTTTCATCTTGTGGGACTTCTTCAT 1114
1178 GCGTCCATGCGGTGGCGGTTTGAAGACTTCTTATCCCGATCGCGAGCTCTCTCT 1237
1115 CGACGCGCTCGCGCGCGCTGTTGCTGGGAGCAGCGCGGAGATAATCTCTCTCT 1174
1238 CGACGCGCTCGGAGCGCGCTCTCTCTCGCGGCGCGGCGGATGGCATCTCTCT 1297
1175 CATCGGCTCGCGCGCGCTCACGCTGTCGCGGAGCGCGCGGAGCGCGCAAGATACCTTG 1234
1298 CTTACCTTCGCGCGCGCTGCTCATGATGACCGCGCGGAGGCGGAGCGGAGCG 1357
1235 GGCATCGGCTTAAGCATCGCTTCAACCTCGCTACGCTGCGCTTCTTCTTCATCGGCT 1294
1358 GCTAGCGCATTAAGCATCGCGCGCTGCTGCTGTTGCGGTGCTTCTCGCTCGGCT 1417
1295 TGGCCCATCAAGTGGGTGTAAGTGGGATCTTCCGCTTCAGGTGCGCGCGCTGCG 1354
1418 GGGCGCGCTCGGCTGAGTACAGCTCGGAGATCTACCGGTGCGGCTGCGTGGCGAGCG 1477
1355 CTGCTGCTCGGCTCGCGCGCAACCGCTCACGAGCGGCTCATCTTCTTCATGACCTTCT 1414
1478 GCGGCGCATCGGAGCGGCTGAACGCGCTGATGAGCGGTGCGCACCATGCTCTCTCT 1537
1415 GTGCTGTCCAAAGGCGCATCAACATCGCGGAGCTTCTTCTTACTCTCGGATCGCGCG 1474
1538 CTTGCTCTCCAAAGCGCATCAACATCGCGGAGCTTCTTACTGTACGCTTCATCGCGCG 1597

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 02:16:29 ; Search time 7788 Seconds
(without alignments)
7733.952 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 2017
Sequence: 1 cttacatgtatgctgcgc.....aaaaaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estm:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hcc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_hcc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	796	39.5	804	14	CB621641
2	764.4	37.9	768	14	CB620394
3	763.2	37.8	824	14	CB621640
4	746	37.0	818	14	CB625172

C	5	737.6	36.6	764	14	CB678733
	6	722.6	35.8	723	14	CB675155
	7	706.8	35.0	942	29	CG374317
	8	705	35.0	710	14	CB678732
C	9	687.4	34.1	697	14	CB620333
C	10	675.2	33.5	1006	29	CG283892
	11	656.4	32.5	671	14	CB620332
	12	654	32.4	700	9	AU057756
C	13	649	32.2	658	14	CB620395
C	14	648.6	32.2	929	29	CG260290
C	15	635	31.5	972	29	CG064683
C	16	625	31.0	641	14	CB648567
C	17	621.8	30.8	853	29	CG189646
C	18	609.8	30.2	735	14	CB675156
C	19	592.2	29.4	597	9	AU097469
	20	580	28.8	714	9	AU163952
	21	576.4	28.6	850	29	CG260280
	22	576	28.6	884	28	CG378810
C	23	575.8	28.5	689	9	AU095961
	24	575.8	28.5	852	29	CG731869
C	25	556	27.6	868	29	CG041546
C	26	554.4	27.5	835	29	CG284028
C	27	552.6	27.4	926	29	CG374307
C	28	538.8	26.7	823	29	CG645436
	29	523.4	25.9	920	28	CG443656
C	30	522.8	25.9	954	29	CG612757
C	31	520.4	25.8	786	29	CG222608
C	32	515.8	25.7	782	14	CB625173
C	33	515.8	25.6	802	29	CG309549
C	34	515.2	25.5	699	12	B1810556
	35	512	25.4	736	28	B2418680
C	36	506	25.1	908	28	CG295842
	37	491.6	24.4	821	10	B264405
	38	489.2	24.3	736	14	CG82653
C	39	476.2	23.6	849	29	CG041551
C	40	475.6	23.6	758	28	B2635288
C	41	467.8	23.2	882	28	CG348900
	42	464.4	23.0	902	29	CG283882
	43	463	23.0	800	14	CB623090
C	44	461.4	22.9	768	29	CG333772
	45	458.4	22.7	932	29	CG170800

ALIGNMENTS

RESULT 1	CB621641/c	804 bp	mRNA	linear	EST 08-APR-2003
LOCUS	OSIIa07K22.r OSIIa Oryza sativa (indica cultivar-group) cDNA				
DEFINITION	clone OSIIa07K22 3', mRNA sequence.				
ACCESSION	CB621641				
VERSION	CB621641.1	GI:29616629			
SOURCE	Oryza sativa (indica cultivar-group)				
ORGANISM	Oryza sativa (indica cultivar-group)				
REFERENCE	Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.				
AUTHORS	Bukavota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.				
TITLE	1 (bases 1 to 804)				
JOURNAL	Unpublished (2003)				
COMMENT	Contact: Rod Wing				
	Arizona Genomics Institute				
	University of Arizona				
	Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ				
	85721-0088, USA				
	Tel: 520 626 3967				
	Fax: 520 621 9288				
	Email: http://genome.arizona.edu				

PCR Primers

FORWARD: gta aac cga cgg cca gtg
 BACKWARD: gga aac agc tat gac cat g
 Plate: 07 row: K column: 22
 Seq primer: gga aac agc tat gac cat g.

FEATURES

source
 1..804
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="IR36"
 /db_xref="taxon:39946"
 /clone="OSIIEA07K22"
 /issue_type="leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSIIEa"
 /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Lesion Mimic SPL 11"

ORIGIN

Query Match 39.5%; Score 796; DB 14; Length 804;
 Best Local Similarity 99.4%; Pred. No. 2.6e-64;
 Matches 799; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1155 GCGGGATATCTCTCCCTCATCGGCTCGGCGCGGCTCACCGTCTCGGCGGAC 1214
 DB 804 GCGGGATATCTCTCCCTCATCGGCTCGGCGCGGCTCACCGTCTCGGCGGAC 745
 QY 1215 CCCGACGCAAGATACCTTGGGCCATCGGCTTAAGCATCGGCTCCACCTCGCCTAGTC 1274
 DB 744 CCCGACGCAAGATACCTTGGGCCATCGGCTTAAGCATCGGCTCCACCTCGCCTAGTC 685
 QY 1275 GCCTTCTTCTCATCGGCTTGGGCCCATCGGCTTAAGCATCGGCTTAAGCATCGGCT 1334
 DB 684 GCCTTCTTCTCATCGGCTTGGGCCCATCGGCTTAAGCATCGGCTTAAGCATCGGCT 625
 QY 1335 CTCAGGTGCGCGCTCGGCTGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1394
 DB 624 CTCAGGTGCGCGCTCGGCTGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 565
 QY 1395 GTCATCTCATGACCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1454
 DB 564 GTCATCTCATGACCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 505
 QY 1455 CTCCTACTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1514
 DB 504 CTCCTACTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 445
 QY 1515 CCGCGCGGACGCTGGAGGATGAGCAAGCTGTTCCGCGACACGCGCGCGCTCGGAA 1574
 DB 444 CCGCGCGGACGCTGGAGGATGAGCAAGCTGTTCCGCGACACGCGCGCGCTCGGAA 385
 QY 1575 TCAGAGAGCCAGCAAG 1634
 DB 384 TCAGAGAGCCAGCAAG 325
 QY 1635 TAACCGCAAAATCAACAAATCTTAAGGCTTTTCTTGAAGAAAGCTGCTGCTGCTGCTGCT 1694
 DB 324 TAACCGCAAAATCAACAAATCTTAAGGCTTTTCTTGAAGAAAGCTGCTGCTGCTGCTGCT 265
 QY 1695 GCTAGCAAGTAGTACGACAG 1754
 DB 264 GCTAGCAAGTAGTACGACAG 205
 QY 1755 CGCGCGGCGACACAAAGCTGAGTCCAGCTCGAGACTTCTTAAATCACTTCAAGTAC 1814
 DB 204 CGCGCGGCGACACAAAGCTGAGTCCAGCTCGAGACTTCTTAAATCACTTCAAGTAC 145
 QY 1815 ATGATTTTATTTGCTCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTT 1874
 DB 144 ATGATTTTATTTTCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTT 85
 QY 1875 CAGTATGTAGCAAGGCTGAGGTTGTGTAGTACTAGTACTAGTACTAGTACTAGTACTAGT 1934

DB 84 CAGTATGTAGCAAGGCTGAGGTTGTGTAGTACTAGTACTAGTACTAGTACTAGTACT 25
 QY 1935 GTAAGAAATGTTTAACTGTTAATT 1958
 DB 24 GTAAGAAATGTTTAACTGTTAATT 1

RESULT 2

CB620394

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

PCR Primers

FORWARD: gta aac agc tat gac cat g

Plate: 05 row: J column: 08

Seq primer: gta aac agc tat gac cat g.

Location/Qualifiers

1..768

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="mRNA"

/cultivar="IR36"

/db_xref="taxon:39946"

/clone="OSIIEA05J08"

/issue_type="leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_lib="OSIIEa"

/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Lesion Mimic SPL 11"

Query Match 37.9%; Score 764.4; DB 14; Length 768;

Best Local Similarity 99.6%; Pred. No. 2.1e-61;

Matches 765; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 41 CGCCACTACTGTACAGCGCCAG 100

DB 1 CGCCACTACTGTACAGCGCCAG 60

QY 101 TTCGCGCGGCTGCGCGAGGCGCTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 160

DB 61 TTCGCGCGGCTGCGCGAGGCGCTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

QY 161 CGCCTGCGCATCTCTCGCTCCATGACCTCCATGACCTCCATGACCTCCATGACCTCCATGACCT 220

DB 121 CGCCTGCGCATCTCTCGCTCCATGACCTCCATGACCTCCATGACCTCCATGACCTCCATGACCT 180

QY 221 GAGCGGCGGCTCGCTGTACATCAAG 280

```
Db 181 GAGCGGGCGCTCGCTGTACATCAAGAAGGACTTCAACATCAGTACGCGGAAGTGGAGGT 240
Qy 281 TCTCATGGGATATAGTAACCTTACTCGCTCATCGCTCTTCCGCGCGGGCGGACGTC 340
Db 241 TCTCATGGGATATAGTAACCTTACTCGCTCATCGCTCTTCCGCGCGGGCGGACGTC 300
Qy 341 GGAATGGATCGGCGGGGATACCAATCGCTGTTCCGCGCGGTCAATATCTTCCGCGGGG 400
Db 301 GGAATGGATCGGCGGGGATACCAATCGCTGTTCCGCGCGGTCAATATCTTCCGCGGGG 360
Qy 401 GTTCCCTCATGGGTTGCGCGTCACTACGCTCATGCTTCGCGCGGTCTTCTGCGCGG 460
Db 361 GTTCCCTCATGGGTTGCGCGTCACTACGCTCATGCTTCGCGCGGTCTTCTGCGCGG 420
Qy 461 CATCGCGGTGGGTACGCGTCACTACGCTTCGCGCGGTTCACACGCGGAGTGTGCGCGG 520
Db 421 CATCGCGGTGGGTACGCGTCACTACGCTTCGCGCGGTTCACACGCGGAGTGTGCGCGG 480
Qy 521 GTCGCGGTGGGTTCGCTGACGCTGTTCCGCGGAGTGTTCATCACTTCGCGCATCTGCT 580
Db 481 GTCGCGGTGGGTTCGCTGACGCTGTTCCGCGGAGTGTTCATCACTTCGCGCATCTGCT 540
Qy 581 CGGTAGCTCTCGAATATGCTTCTCCGCTTTCGCGGAGTGTTCATCACTTCGCGCATCT 640
Db 541 CGGTAGCTCTCGAATATGCTTCTCCGCTTTCGCGGAGTGTTCATCACTTCGCGCATCT 600
Qy 641 GTCGCGGTGGGTTCGCTGACGCTGTTCCGCGGAGTGTTCATCACTTCGCGCATCTGCT 700
Db 601 GTCGCGGTGGGTTCGCTGACGCTGTTCCGCGGAGTGTTCATCACTTCGCGCATCTGCT 660
Qy 701 GTCGCGGTGGGTTCGCTGACGCTGTTCCGCGGAGTGTTCATCACTTCGCGCATCTGCT 760
Db 661 GTCGCGGTGGGTTCGCTGACGCTGTTCCGCGGAGTGTTCATCACTTCGCGCATCTGCT 720
Qy 761 GACCTCCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGAGTGTTCATCACTTCGCGCAT 808
Db 721 GACCTCCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGAGTGTTCATCACTTCGCGCAT 768

RESULT 3
CB621640
LOCUS CB621640
DEFINITION OSIIIEa07K22.f OSIIIEa Oryza sativa (indica cultivar-group) cDNA.
ACCESSION CB621640
VERSION CB621640.1 GI:29616628
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 824)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
JOURNAL between rice and Magnaporthe grisea
COMMENT Unpublished (2003)
CONTACT: Rod Wing
INSTITUTION Arizona Genomics Institute
UNIVERSITY University of Arizona
ADDRESS Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
CITY 85721-0088, USA
TEL: 520 626 3967
FAX: 520 621 9288
EMAIL: http://genome.arizona.edu
PCR PRIMERS
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 07 row: K column: 22
Seq primer: gta aaa cga cgg cca gtc.
FEATURES
source Location/Qualifiers
1..824
```

```

/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mrna"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIIEa07K22"
/tissue_type="Leaf"
/dev_stages="3 week"
/lab_host="DH10B"
/clone_lib="OSIIIEa"
/notes="Vector: pbluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Lesion Mimic SPL 11"

ORIGIN
Query Match 37.8%; Score 763.2; DB 14; Length 824;
Best Local Similarity 95.4%; Pred. No. 2.6e-61;
Matches 786; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
874 GAAACGAGAAGCGGTGTGGAAGAGGACTCATCTTCCCGGACCCCGGCCATCGCGCGCA 933
Db 1 GAAACGAGAAGCGGTGTGGAAGAGGACTCATCTGTCCCGACCCCGGCCATCGCGCGCA 60
Qy 934 TCTCTGTGTCGGGATCGGCATTCATCTTCCAGCATGCGTTGGGCAATTCATCCGTCG 993
Db 61 TCTCTGTGTCGGGATCGGCATTCATCTTCCAGCATGCGTTGGGCAATTCATCCGTCG 120
Qy 994 TCTTCTACAGCCCTCTCGTGTTCAGAGCCCGGATTAAAGCAACGACAAACACTCTTTGG 1053
Db 121 TCTCTACAGCCCGCGGTGTTCAGAGCGCGGGATCACGACGACAAACACTCTTCG 180
Qy 1054 GCACCACTTGGCCGTTCTGTTGTACCAAGAGGCTTTTCATCTTGTGCGACTTTCTTCA 1113
Db 181 GCACCACTTGGCCGTTCTGTTGTACCAAGAGGCTTTTCATCTTGTGCGACTTTCTTCC 240
Qy 1114 TCGAGGGGTGTCGGCGCGCGCGCTGTTCTGTCGGGACGACGGGCGGATATCTCTCCC 1173
Db 241 TCGACCGGTGTCGGCGCGCGCGCTGCTGCTGAGCAGCAGCGGCGGATGATCTCTCTCC 300
Qy 1174 TCATCGGCTCGCGCGCGCGGTACCGTCTGTCGGCAGCACCCCGACGCAAGATACCTT 1233
Db 301 TCATCGGCTCGCGCGCGCGGTACCGTCTGTCGGCAGCACCCCGACGCAAGATACCTT 360
Qy 1234 GGGCATCGGCTAAGCATTCGCTTCCAGCTGCTGCTAGTGCCTTTCTTCCATCGGCG 1293
Db 361 GGGCATCGGCTAAGCATTCGCTTCCAGCTGCTGCTAGTGCCTTTCTTCCATCGGCG 420
Qy 1294 TTGGCCCATCAGTGGGTGTACAGCTCGAGATCTTCCGCTCCAGTTCGCGCGCTGG 1353
Db 421 TTGGCCCATCAGTGGGTGTACAGCTCGAGATCTTCCGCTCCAGTTCGCGCGCTGG 480
Qy 1354 GCTGCTCGCTCGCGCTCGCGCAACCGCGTCACCGCGGCTCATCTCCATGACCTTCC 1413
Db 481 GCTGCTCGCTCGCGCTCGCGCAACCGCGTCACCGCGGCTCATCTCCATGACCTTCC 540
Qy 1414 TGTGCTGTCGAAGGCTACCATTCGCGGAGCTTCTTCTTACTCCGCGATCGCG 1473
Db 541 TGTGCTGTCGAAGGCTACCATTCGCGGAGCTTCTTCTTACTCCGCGATCGCG 600
Qy 1474 CGCTCGCTCGGTGTTCTTCTACCTACCTCCCGAGACCCCGCGCGCGAGCGCTGGAG 1533
Db 601 CGCTCGCTCGGTGTTCTTCTACCTACCTCCCGAGACCCCGCGCGCGAGCGCTGGAG 660
Qy 1534 AGATGAGCAAGTGTTCGCGGACGACGCGCGCTCGGAATTCAGACGAGCGCGCAAG 1593
Db 661 AGATGAGCAAGTGTTCGCGGACGACGCGCGCTCGGAATTCAGACGAGCGCGCAAG 720
Qy 1594 AGAAGCAAGTGTTCGCGGACGACGCGCGCTCGGAATTCAGACGAGCGCGCAAG 1653
Db 721 AGAAGCAAGTGTTCGCGGACGACGCGCGCTCGGAATTCAGACGAGCGCGCAAG 780
Qy 1654 TCTAAGGGTCTTCTTGCAAAAACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1697
Db 781 TCTAAGGGTCTTCTTGCAAAAACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824
```

```

RESULT 4
CB625172
LOCUS
DEFINITION
  CB25172.f OSIIIEA Oryza sativa (indica cultivar-group) cDNA
  clone OSIIIEA14C12 5', mRNA sequence.
ACCESSION
  CB625172
VERSION
  CB625172.1 GI:29620160
KEYWORDS
  EST.
SOURCE
  Oryza sativa (indica cultivar-group)
  Oryza sativa (indica cultivar-group)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzeae; Oryza.
  1 (bases 1 to 818)
REFERENCE
  Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
  Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  Large-scale identification of ESTs involved in the interaction
  between rice and Magnaporthe grisea
  Unpublished (2003)
JOURNAL
  Contact: Rod Wing
  Arizona Genomics Institute
  University of Arizona
  Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
  85721-0088, USA
  Tel: 520 626 3967
  Fax: 520 621 9288
  Email: http://genome.arizona.edu
PCR PRIMERS
  FORWARD: gta aaa cga cgg cca gtg
  BACKWARD: gga aac agc tat gac cat g
  Plate: 14 row: C column: 12
  Seq primer: gta aaa cga cgg cca gtg.
  Location/Qualifiers
  1..818
  /organism="Oryza sativa (indica cultivar-group)"
  /mol_type="mRNA"
  /cultivar="IR36"
  /db_xref="taxon:39946"
  /clone="OSIIIEA14C12"
  /tissue type="Leaf"
  /dev stage="3 week"
  /lab_host="DH10B"
  /clone_lib="OSIIIEA"
  /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
  XhoI; Lesion Mimic SPL 11"
ORIGIN
  Query Match 37.0%; Score 746; DB 14; Length 818;
  Best Local Similarity 94.5%; Pred No. 9.7e-60;
  Matches 773; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
  QY 457 CCGCATCGGCGTGGCTACCGCTCATGATCGCGCGGTGTACACCGCGAGGTGTCG 516
  Db 1 CCGCATCGGCGTGGCGCACTCGCTCATGATCGCGCGGTGTACACCGCGAGGTGTCG 60
  QY 517 CGCGTCGGCGTGGCTTCTGACGTCGTTCCGGAGGTGTTTCATCACTTCGGCATCC 576
  Db 61 CGCGTCGGCGTGGCTTCTGACGTCGTTCCGGAGGTGTTTCATCACTTCGGCATCC 120
  QY 577 TGCTCGGTCAGCTTCGAACATATGCTTTCTCCCGCTTGCCTGAACTCGGTTGGCGCA 636
  Db 121 TGCTCGGTCAGCTTCGAACATATGCTTTCTCCCGCTTGCCTGAACTCGGTTGGCGCA 180
  QY 637 TCATGCTCGGTCAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 696
  Db 181 TCATGCTCGGTCAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
  QY 697 CGGAGTCGCGCGTGGTGGTGTATGAAGGACGCTCGCGGACGCAAGTGTGTGCTGG 756
  Db 241 CGGAGTCGCGCGTGGTGGTGTATGAAGGACGCTCGCGGACGCAAGTGTGTGCTGG 300
  QY 757 AGAAGACCTCCGACACGCGGAGGCGCGCGGAGCGCTTGGCCGACATCAAGCGCGCG 816

```

```

Db 301 AGAAGACCTCCGACACGCGGAGGCGCGGAGGCGCTGCGCCACATCAAGCGCGCG 360
QY 817 CCGGCATCCCTAGAGAGCTCGACGCGAGCTGTCACCGTCCCAAGAGAGGAGCGGAA 876
Db 361 CCGGCATCCCGAGGAGCTCGACGCGAGCTGTCACCGTCCCAAGAGAGGAGCGGAA 420
QY 877 ACAGAAGCGGCTGTGGAAGGAGCTCATCTCTCCCGACCGCCGATCGGCGCATCC 936
Db 421 ACAGAAGCGGCTGTGGAAGGAGCTCATCTCTCCCGACCGCCGATCGGCGCATCC 480
QY 937 TGCTGTCCGGATTCGATCCACTTCTTCCAGCATGCGTTGGGCAATTCATCTCCGTCGT 996
Db 481 TGCTGTCCGGATTCGATCCACTTCTTCCAGCATGCGTTGGGCAATTCATCTCCGTCGT 540
QY 997 TCTACAGCCCTCTGCTGTTCAGAGCGCCGCGGATTAACGACGACAAACACTTCTTGGCA 1056
Db 541 TCTACAGCCCTCTGCTGTTCAGAGCGCCGCGGATTAACGACGACAAACACTTCTTGGCA 600
QY 1057 CCACCTTGGCGCTTGGCTGTTCACCAAGAGGCTTTTCATCTTGTGGCGACTTTCTTCATCG 1116
Db 601 CCACCTTGGCGCTTGGCTGTTCACCAAGAGGCTTTTCATCTTGTGGCGACTTTCTTCATCG 660
QY 1117 ACCGCGTCCGGCGCGCGCTGCTTCTGGGAGCAGCGCGGAGTATCTTCTTCTCTCA 1176
Db 661 ACCGCGTCCGGCGCGCGCTGCTTCTGGGAGCAGCGCGGAGTATCTTCTTCTCTCA 720
QY 1177 TCGGCGCTCCGGCGCGCGCTGCTTCTGGGAGCAGCGCGGAGTATCTTCTTCTCTCA 1236
Db 721 TCGGCGCTCCGGCGCGCGCTGCTTCTGGGAGCAGCGCGGAGTATCTTCTTCTCTCA 780
QY 1237 CCATCGGCTTAAGCATGCGCTTCCACCCCTCGCTACGTC 1274
Db 781 CCATCGGCTTAAGCATGCGCTTCCACCCCTCGCTACGTC 818

RESULT 5
CB678733/c
LOCUS
DEFINITION
  CB678733 OSUNEF Oryza sativa (japonica cultivar-group) cDNA
  clone OSUNEF01D24 3', mRNA sequence.
ACCESSION
  CB678733
VERSION
  CB678733.1 GI:29682458
KEYWORDS
  EST.
SOURCE
  Oryza sativa (japonica cultivar-group)
  Oryza sativa (japonica cultivar-group)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzeae; Oryza.
  1 (bases 1 to 764)
REFERENCE
  Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
  Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  Large-scale identification of ESTs involved in the interaction
  between rice and Magnaporthe grisea
  Unpublished (2003)
JOURNAL
  Contact: Rod Wing
  Arizona Genomics Institute
  University of Arizona
  Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
  85721-0088, USA
  Tel: 520 626 3967
  Fax: 520 621 9288
  Email: http://genome.arizona.edu
PCR PRIMERS
  FORWARD: gta aaa cga cgg cca gtg
  BACKWARD: gga aac agc tat gac cat g
  Plate: 01 row: D column: 24
  Seq primer: gga aac agc tat gac cat g.
  Location/Qualifiers
  1..764
  /organism="Oryza sativa (japonica cultivar-group)"
  /mol_type="mRNA"
  /cultivar="Nipponbare"

```



```
ORIGIN
Query Match 36.6%; Score 737.6; DB 14; Length 764;
Best Local Similarity 98.6%; Pred. No. 5.9e-59;
Matches 755; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 1219 ACSCCAAGATACCTTCGGGCATCGGCTAAGATCGCTCCACCTCGCTACGCTCGCT 1278
DB 764 ACSCCAAGATACCTTCGGGCATCGGCTAAGATCGCTCCACCTCGCTACGCTCGCT 705

QY 1279 TCTTCTCCATCGGCTTCGCCCATCACGTGGGTGTACAGTCCGAGATCTTCCCGCTCC 1338
DB 704 TCTTCTCCATCGGCTTCGCCCATCACGTGGGTGTACAGTCCGAGATCTTCCCGCTCC 645

QY 1339 AGTGCGCGCTGGCTGCTGCTCGGCTCGCCGCAACCGCTCACCGGCGTCA 1398
DB 644 AGTGCGCGCTGGCTGCTGCTCGGCTCGCCGCAACCGCTCACCGGCGTCA 585

QY 1399 TCTCCATGACCTTCCTGCTGCTCCAAAGGCATCACCATCGGCGGAGCTTCTTCTCT 1458
DB 584 TCTCCATGACCTTCCTGCTGCTCCAAAGGCATCACCATCGGCGGAGCTTCTTCTCT 525

QY 1459 ACTCCGGCATCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1518
DB 524 ACTCCGGCATCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 465

QY 1519 GCGGAGCTGGAGGAGATGAGCAAGCTGTTCGGCAGCGGCGGCTCGGAATCAG 1578
DB 464 GCGGAGCTGGAGGAGATGAGCAAGCTGTTCGGCAGCGGCGGCTCGGAATCAG 405

QY 1579 ACAGCGCAGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1638
DB 404 ACAGCGCAGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345

QY 1639 CGCAAAATACCAAAATCCTTAAGGTTTCTTTCGCAAAAGAGTGCTGCTGCTGCTA 1698
DB 344 CGCAAAATACCAAAATCCTTAAGGTTTCTTTCGCAAAAGAGTGCTGCTGCTGCTA 285

QY 1699 GCAAGTAGTAGCAACAGCTGGGAAGATTCGCTGATCCGCGGCTGCTGGAGAGCGAG 1758
DB 284 GCAAGTAGTAGCAACAGCTGGGAAGATTCGCTGATCCGCGGCTGCTGGAGAGCGAG 225

QY 1759 CGCGGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAATCATCTTCAAGTACATGG 1818
DB 224 CGCGGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAATCATCTTCAAGTACATGG 165

QY 1819 ATTTTATTTTTCCTTTGCTTTCGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1878
DB 164 ATTTTATTTTTCCTTTGCTTTCGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 105

QY 1879 ATGTAGCAAGCTGAGGTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1938
DB 104 ATGTAGCAAGCTGAGGTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 45

QY 1939 GAAATGTTTAACTGTTTAACTAAGCAGTATTTGTCAGTAACTAAA 1984
DB 44 GAAATGTTTAACTGTTTAACTAAGCAGTATTTGTCAGTAACTAAA 1

RESULT 6
CB675155 OSJNE10K23.f OSJNE Oryza sativa (japonica cultivar-group) cDNA
LOCUS
DEFINITION
clone OSJNE10K23 5', mRNA sequence.
ACCESSION
CB675155
```

```
VERSION CB675155.1 GI:29678880
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 723)
AUTHORS Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
TITLE Large-scale identification of ESTs involved in the interaction
JOURNAL between rice and Magnaporthe grisea
COMMENT Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 10 row: K column: 23
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1. 723
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNE10K23"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNE"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"
ORIGIN
Query Match 35.8%; Score 722.6; DB 14; Length 723;
Best Local Similarity 99.9%; Pred. No. 1.4e-57;
Matches 722; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 35 CTCGACCGCCACTACTGTACACGCGCCAGAGAGCTCTCTCTCTCTGACACACCGGA 94
DB 1 CTCGACCGCCACTACTGTACACGCGCCAGAGAGCTCTCTCTCTCTGACACACCGGA 60

QY 95 GATGCTTTCGCGCGCTGCGGAGGCGCTCGCGCGAAGAGAGGCAACGTCGGGT 154
DB 61 GATGCTTTCGCGCGCTGCGGAGGCGCTCGCGCGAAGAGAGGCAACGTCGGGT 120

QY 155 CGCTTCGCTCGCGCATCTTCGCTCCATCCATCCATCCATCCATCCATCCATCCATCCG 214
DB 121 CGCTTCGCTCGCGCATCTTCGCTCCATCCATCCATCCATCCATCCATCCATCCG 180

QY 215 GGTGATGAGCGGGGCGCTGCTGATCAACAAGAGCTTCAACATCAGTACGCGAGGT 274
DB 181 GGTGATGAGCGGGGCGCTGCTGATCAACAAGAGCTTCAACATCAGTACGCGAGGT 240

QY 275 GGAGGTTCTCATGGGCATCTGAACTCTACTCGCTCATCGGCTCTCTCGCGGGGCGG 334
DB 241 GGAGGTTCTCATGGGCATCTGAACTCTACTCGCTCATCGGCTCTCTCGCGGGGCGG 300

QY 335 GACGTGCGACTGATCGCGCGGCGGTACACCATGTTGTCGCGCGGCTCATATTTTCG 394
DB 301 GACGTGCGACTGATCGCGCGGCGGTACACCATGTTGTCGCGCGGCTCATATTTTCG 360

QY 395 GGGGSGGTTCTCATGGGTTTCGCGGTCACATAGCCATGCTCATGTTTCGCGCGGCTTCG 454
DB 361 GGGGSGGTTCTCATGGGTTTCGCGGTCACATAGCCATGCTCATGTTTCGCGCGGCTTCG 420
```

QY 455 GCGCGCATCGCGGTGGGCTACGGCTCATGATCGCGCGGTGTACACCGCGAGGTGTC 514
 Db 421 GCGCGCATCGCGGTGGGCTACGGCTCATGATCGCGCGGTGTACACCGCGAGGTGTC 480
 QY 515 GCGCGGTGCGCGGTGGGCTTCCTGAGTGTGTTCCCGGAGGTTCATCAACTTCGGCAT 574
 Db 481 GCGCGGTGCGCGGTGGGCTTCCTGAGTGTGTTCCCGGAGGTTCATCAACTTCGGCAT 540
 QY 575 CTTGCTCGGTACGTCCTCGAACTATGTTTCTCCCGGTTGCGGCTGAACCTCGGTGGCG 634
 Db 541 CTTGCTCGGTACGTCCTCGAACTATGTTTCTCCCGGTTGCGGCTGAACCTCGGTGGCG 600
 QY 635 CATCATGCTCGGCATCGCGCGCGCGCTCGTCTGCTCGCGCTCATGTTGTCGCGCAT 694
 Db 601 CATCATGCTCGGCATCGCGCGCGCGCTCGTCTGCTCGCGCTCATGTTGTCGCGCAT 660
 QY 695 GCGGAGTCGCGCGGTGGGCTCATGAAGGAGCGCTCGCGACGCCAAGGTGGTGT 754
 Db 661 GCGGAGTCGCGCGGTGGGCTCATGAAGGAGCGCTCGCGACGCCAAGGTGGTGT 720
 QY 755 GGA 757
 Db 721 GGA 723

RESULT 7
 CG374317 942 bp DNA linear GSS 26-AUG-2003
 LOCUS OGYB078TV ZM 0.7 1.5 KB Zea mays genomic clone ZM5Bma0642M11,
 DEFINITION genomic survey sequence.

ACCESSION CG374317
 VERSION CG374317.1 GI:34291584
 KEYWORDS GSS.

SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

AUTHORS Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.

TITLE Consortium for Maize Genomics
 JOURNAL Unpublished (2002)
 COMMENT Other GSSs: OGYB078TH
 Contact: Cathy Whitelaw

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..942

/organism="Zea mays"

/mol type="genomic DNA"

/strain="B73"

/db xref="taxon:4577"

/clone lib="ZM 0.7 1.5 KB"

/note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

ORIGIN

Query Match 35.0%; Score 706.8; DB 29; Length 942;

Best Local Similarity 84.4%; Pred. No. 3.6e-56;

Matches 795; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 434 GCTCATGTTCCGGCGCTTCGTCGGCGGATCGCGGTGGGCTACGCGCTCATGATCGCGC 493

Db 1 GCTCATGTTCCGGCGCTTCGTCGGCGGATCGCGGTGGGCTACGCGCTCATGATCGCGC 60

QY 494 GGTGTACACCGCGAGGTGTCCCGCGGTTCGGCGGTTCCTGACGTCTGTTCCCGGA 553
 Db 61 CGTGTACACCGCGAGGTGTCCCGCGGTTCGGCGGTTCCTGACGTCTGTTCCCGGA 120
 QY 554 GGTGTTCATCAACTTCGGCATCTCTCGGTAGTCTCGAACTATGTTTCTCCCGTT 613
 Db 121 GGTGTTCATCAACTTCGGCATCTCTCGGTAGTCTCGAACTATGTTTCTCCCGTT 180
 QY 614 GCGCGTGAACCTTCGGGTGGCGCATCATGCTCGGCATCGCGCGGCGCGCTCGTGTCT 673
 Db 181 CCGGTTCGGGTGGCGCATCATGCTCGGCATCGCGCGGCGCGCTCGTGTCT 240
 QY 674 CCGGTTCATGTTGCTTCGGCATCGCGGAGTTCGCGGTGGTGGTTCATGAAGGAGCGCT 733
 Db 241 CCGGTTCATGTTGCTTCGGCATCGCGGAGTTCGCGGTGGTGGTTCATGAAGGAGCGCT 300
 QY 734 CCGCGACGCCAAGGTGGTCTGGAGAACCTTCGACACCGCGGAGGAGCGCGGAGCG 793
 Db 301 GCGGACGCCAAGGTGGTCTGGAGAACCTTCGACACCGCGGAGGAGCGCGGAGGAG 360
 QY 794 CTTGCGGCATCAAGCGCGCGCGCATCCCTGAGGAGTTCGACCGCGAGCTGGTGTAC 853
 Db 361 GCTGCTGACATCAAGCGCGCGCGCATCCCGCAGGAGTTCGACCGCGAGCTGGTGTAC 420
 QY 854 CTTGCGGCATCAAGCGCGCGCGCATCCCTGAGGAGTTCGACCGCGAGCTGGTGTAC 913
 Db 421 CTTGCGGCATCAAGCGCGCGCGCATCCCTGAGGAGTTCGACCGCGAGCTGGTGTAC 480
 QY 914 GACCCCGGCGCATCGCGCGCATCTCTCGCGGATCGGCATCCACATTCCTCCAGCATGC 973
 Db 481 GACCCCGGCGCATCGCGCGCATCTCTCGCGGATCGGCATCCACATTCCTCCAGCATGC 540
 QY 974 GTTGGGATTCATCTCGTCTCTTCTCAGAGCTCTCTGTTTCAAGAGCCCGCGATTAAC 1033
 Db 541 GTTGGGATTCATCTCGTCTCTTCTCAGAGCTCTCTGTTTCAAGAGCCCGCGATTAAC 600
 QY 1034 GAAACACAAAACATCTTTGGGACACCATTTGGCGTTCGGTGTCCACCAAGAGGCTTTTCAT 1093
 Db 601 CAGCGACAAACAGATCTCTGGGAAAGCATCTGGCGGTGGCGTTCACCAAGAGCTTTTCAT 660
 QY 1094 CTTGTTGGGATCTTTCTTCATCGACGGGTTCGGCGCGCGCGTGTCTGGGAGCAGC 1153
 Db 661 CTTGTTGGGATCTTTCTTCATCGACGGGTTCGGCGCGCGCGTGTCTGGGAGCAGC 720
 QY 1154 GGGCGGATATCTCTCCCTCATCGGCTTCGGCGCGCGGCTCACCGTCTCGCGGAGCA 1213
 Db 721 GGGCGGATATCTCTCCCTCATCGGCTTCGGCGCGCGGCTCACCGTCTCGCGGAGCA 780
 QY 1214 CCGCGACGCCAAGATACCTTGGGCGCATCGGCTTAAGCATCGCCTCCACCTTCGCTACGT 1273
 Db 781 CCAGGACGAAAAGATCCCATGGGCGCATCGGCGTGGCGTCTACCATGGCTACGT 840
 QY 1274 GCGCTTCTTCCATCGGCGCTTGGCGCGCATCACGTTGGGTGTACAGCTCGGAGATCTTCCC 1333
 Db 841 GCGCTTCTTCCATCGGCGCTTGGCGCGCATCACGTTGGGTGTACAGCTCGGAGATCTTCCC 900
 QY 1334 GCTCAGGTTCGGCGCGCTGGGCTGCTGCTCGGCGTGGCGCG 1375
 Db 901 GCTCATGTTCCGGCGGATGGGCTGGCGGTGGCGCTGGCGCTC 942

RESULT 8

CB678732

LOCUS

DEFINITION

OSUNef01D24.f OSUNef Oryza sativa (japonica cultivar-group) cDNA

clone OSUNef01D24 5', mRNA sequence.

ACCESSION CB678732

VERSION CB678732.1

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

CB678732 710 bp mRNA linear EST 09-APR-2003

OSUNef01D24.f OSUNef Oryza sativa (japonica cultivar-group) cDNA

clone OSUNef01D24 5', mRNA sequence.

ACCESSION CB678732

VERSION CB678732.1

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Ehrhartoideae; Oryzae; Oryza.
1 (bases 1 to 710)
Jantassuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: D column: 24
Seq primer: gta aaa cga cgg cca gtc.

FEATURES
source
1..710
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEF01D24"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEF"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"

ORIGIN
Query Match 35.0%; Score 705; DB 14; Length 710;
Best Local Similarity 99.7%; Pred. No. 5.9e-56;
Matches 705; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 33 CACCTGACCGCCACTACTGTACAGCGCCAGAGGAGCGCTCTCTCTCTGTGACACCG 92
DB 4 CTCTGACCGCCACTACTGTACAGCGCCAGAGGAGCGCTCTCTCTGTGACACCG 63
QY 93 GAGATGGTTCGCGCGCTCGGAGCGCTCGGAGCGCTCGGAGCGCGGAGCGGAGCGG 152
DB 64 GAGATGGTTCGCGCGCTCGGAGCGCGCTCGGAGCGCGGAGCGGAGCGGAGCGG 123
QY 153 TTGCGCTTCGCTGCGCCATCTCGCTCCATGACCTCCATCTCTCTCTCTCTCTCTCT 212
DB 124 TTGCGCTTCGCTGCGCCATCTCTGCTTCCATGACCTCTCTCTCTCTCTCTCTCT 183
QY 213 GGGGTGATGAGCGGCGCTGCTGTATCATCAAGAGGACTTCAACATCATGTCGGAAG 272
DB 184 GGGGTGATGAGCGGCGCTGCTGTATCATCAAGAGGACTTCAACATCATGTCGGAAG 243
QY 273 GTGGAGGTTCTCATGGGCTACTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 332
DB 244 GTGGAGGTTCTCATGGGCTACTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 303
QY 333 CGAGCTCGGACTGATCGCGCGCGGTACACCATCTGTTTCGCGCGCGCTCATATTCTTC 392
DB 304 CGAGCTCGGACTGATCGCGCGCGGTACACCATCTGTTTCGCGCGCGCTCATATTCTTC 363
QY 393 GCGGGGGGTTCTCATGGGTTTCCGCTGCTCACTGACCTGCTCATGTTTCGCGCGCTTC 452
DB 364 GCGGGGGGTTCTCATGGGTTTCCGCTGCTCACTGACCTGCTCATGTTTCGCGCGCTTC 423
QY 453 GTGGCGGCTCGGCGGTGAGCGGTATGATCGCGCGGTGATCACCGCGGAGGTG 512
DB 424 GTGGCGGCTCGGCGGTGAGCGGTATGATCGCGCGGTGATCACCGCGGAGGTG 483
QY 513 TCGCGGCGTCCGCGCGTGGCTTCTGACGCTGTTCCGAGGTTTTCATCACTTCGGC 572

DB 484 TCGCGCGCGTTCGGCGCGTGGCTTCTGACCTGTTCTCCCGAGGTGTTTCATCACTTCGCG 543
QY 573 ATCTCTCTCGGTACGTCTCGAATATGTTTCTCCCGCTTCCCGCTGAACTTCGCGTGG 632
DB 544 ATCTCTCTCGGTACGTCTCGAATATGTTTCTCCCGCTTCCCGCTGAACTTCGCGTGG 603
QY 633 CGCATCATGCTCGGCATCGCGCGCGCTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 692
DB 604 CGCATCATGCTCGGCATCGCGCGCGCTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 663
QY 693 ATGCGCGAGTTCGCGCGCGGTGGTGTATCAAGGAGCGCTTCGCGA 739
DB 664 ATGCGCGAGTTCGCGCGCGGTGGTGTATCAAGGAGCGCTTCGCGA 710

RESULT 9
CB620333/c
LOCUS
DEFINITION
clone OSIIeA05H07 3', mRNA sequence.
CB620333
ACCESSION
VERSION
CB620333.1 GI:29615321
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.
REFERENCE
AUTHORS
Jantassuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL
COMMENT
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 05 row: H column: 07
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1..697
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIeA05H07"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIeA"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Lesion Mimic SPL 11"

ORIGIN
Query Match 34.1%; Score 687.4; DB 14; Length 697;
Best Local Similarity 99.1%; Pred. No. 2.4e-54;
Matches 691; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1234 GGGCCATCGGCTTAAGCATCGCTCCACCTCGCTTCTCTCTCTCTCTCTCTCTCTCTCT 1293
DB 697 GGGCCATCGGCTTAAGCATCGCTCCACCTCGCATACGCTCTCTCTCTCTCTCTCTCT 638
QY 1294 TTGCCCCCATCAGTGGGTGATGATCGAGATCTTCCGCTTCCAGGTGCGCGCTGG 1353
DB 637 TTGCCCCCATCAGTGGGTGATGATCGAGATCTTCCGCTTCCAGGTGCGCGCTGG 578

1354 GGTGCTCGCTCGGCTCGCGCCGCAACCGGTCACACGCGGTACACGCGGGGTCTATCTCCATGACCTTCC 1413
 Db GGTGCTCGCTCGGCTCGCGCCGCAACCGGTCACACGCGGTACACGCGGGGTCTATCTCCATGACCTTCC 518
 QY TGTGCTGTCGCTCGGCTCGCGCCGCAACCGGTCACACGCGGTACACGCGGGGTCTATCTCCATGACCTTCC 1473
 Db TGTGCTGTCGCTCGGCTCGCGCCGCAACCGGTCACACGCGGTACACGCGGGGTCTATCTCCATGACCTTCC 458
 QY CGCTCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1474
 Db CGCTCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 398
 QY AGATGAGCAAGCTGTTCGCGGACACGCGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1534
 Db AGATGAGCAAGCTGTTCGCGGACACGCGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 338
 QY AGAAGAAGAGTGGAAATGGCGGCACTAACTGATCAAACTAACTGATCAAACTAACTGATCAAACTAACT 1594
 Db AGAAGAAGAGTGGAAATGGCGGCACTAACTGATCAAACTAACTGATCAAACTAACTGATCAAACTAACT 278
 QY TCTAAGGCTTCTTGTGCAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1654
 Db TCTAAGGCTTCTTGTGCAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 218
 QY AACGTGGGAAGATTTCGCTGATCCGCGGTTGCTGGAGACGACGCGCGGCGGACGCAAAAGC 1714
 Db AACGTGGGAAGATTTCGCTGATCCGCGGTTGCTGGAGACGACGCGCGGCGGACGCAAAAGC 1773
 QY AACGTGGGAAGATTTCGCTGATCCGCGGTTGCTGGAGACGACGCGCGGCGGACGCAAAAGC 158
 Db AACGTGGGAAGATTTCGCTGATCCGCGGTTGCTGGAGACGACGCGCGGCGGACGCAAAAGC 1833
 QY TGAGCTCCAGCTCGAGACTTCTTAAATCATCTTCAAGTATACGATGATGATGATGATGATGATGATGAT 1774
 Db TGAGCTCCAGCTCGAGACTTCTTAAATCATCTTCAAGTATACGATGATGATGATGATGATGATGATGAT 98
 QY TTGCTTTGCTTAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1834
 Db TTGCTTTGCTTAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 38
 QY GGTGCT 1894
 Db GGTGCT 1930
 37 GGTGCT 1

RESULT 10
 LOCUS CG283892/c
 DEFINITION OG1AC79TV ZM 0.7.1.5 KB Zea mays genomic clone ZMMBma0715N13, linear GSS 25-AUG-2003
 genomic survey sequence.
 ACCESSION CG283892
 VERSION CG283892.1
 KEYWORDS GI:34198106
 SOURCE GSS.
 ORGANISM Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1. (bases 1 to 1006)
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OG1AC79TH
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: n/a
 Class: sheared ends.
 Location/Qualifiers
 1..1006
 /organism="Zea mays"

/mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBma0715N13"
 /clone_lib="ZM 0.7.1.5 KB"
 /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 33.5%; Score 675.2; DB 29; Length 1006;
 Best Local Similarity 80.2%; Pred. No. 2.7e-53;
 Matches 804; Conservative 1; Mismatches 194; Indels 3; Gaps 1;

QY 205 ACATATCGGGTGTATGAGGGGGCTGCTGCTACATCAAGAGGACTTCAACATCAGTG 264
 Db 1002 AAGACATCGGGTGTATGAGGGGGCTGCTGCTACATCAAGAGGACTTCAACATCAGTG 943
 QY 265 ACAGGAAGTGGAGGTTCCTCATGGGCATCTCAACCTCTACTCGCTCATCGGCTCCTTCG 324
 Db 942 ACCTGCAGCTGGAGATCCTCATGGGCATCTCAACCTCTACTCGCTCATCGGCTCCTTCG 883
 QY 325 CGCGGGGGGAGAGCTCGGATCGGATCGCGGGGGTACACCATCGTGTTCGCGGGCGCTCA 384
 Db 882 CGCGGGGGGAGAGCTCGGATCGGATCGGTCGCGGGTTCCTGCTGCTGCTGCTGCTGCTGCT 823
 QY 385 TATTCTTCGCGGGGGGTTCCTCATGGGGTTCGCGGTCAACTACGCACTGCTCATGTTCCG 444
 Db 822 TATTCTTCGCGGGGGTTCCTCATGGGGTTCGCGGTCAACTACGCACTGCTCATGTTCCG 763
 QY 445 GCCTCTTCGTGGCCGCATCGCGGTGGGTACGCGCTCATGATTCGCGCGCTGTCACCG 504
 Db 762 GCCTCTTCGTGGCCGCATCGCGGTGGGTACGCGCTCATGATTCGCGCGCTGTCACCG 703
 QY 505 CGAGGTGTCGCGGGGTTCGCGGTGGGTTCCTGACGCTTCCTCGGAGGTTCATCA 564
 Db 702 CGAGGTGTCGCGGGGTTCGCGGTGGGTTCCTGACGCTTCCTCGGAGGTTCATCA 643
 QY 565 ACTTCGGCATCTGCTCGGGTACGCTCGAATATGCTTTCCTCGGCTTCGCGCTGAACC 624
 Db 642 ACCTCGGATCTCTCTCGGCTACGCTTCCTCACTACGCTTCGCGCTTCGCGCTCGGC 583
 QY 625 TCGGTTGGGCATCATGCTTCGCGATCGCGCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 584
 Db 582 TCGGTTGGGCATCATGCTTCGCGATCGGAGCGCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCT 523
 QY 685 TGCTCGGCATGCGCGAGTTCGCGGTGCTGCTCATGAAGGAGCGCTTCGCGGACGCCA 744
 Db 522 TGTTTCGCATGCGCGAGTTCGCGGTGCTGCTCATGAAGGAGCGCTTCGCGGACGCCA 463
 QY 745 AGTGTGCTTGAGAGAACCTTCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGACGCCA 804
 Db 462 GGTGCTGCTGAGAGAACCTTCGACACGCGGAGGAGCGCGGAGCGCTTCGCGGACGCCA 403
 QY 805 TCAAGCGCGCGCGGCATCCCTCC---TGAGGAGCTCGACGCGGACGCTGCTGCTGCTGCTGCTGCT 861
 Db 402 TCAAGCGCGCGCGGCATCCCTCCCGGAGGAGCGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCT 343
 QY 862 AGAGGGGAGCGGAGAACGAGAGCGGGGTGGAAGAGGCTCATCTGCTTCGCGGACGCCG 921
 Db 342 AGAGGAAGAGCGCGGAGGAGCGGGGTGGAAGAGGCTCATCTGCTTCGCGGACGCCG 283
 QY 922 CCATCGCGGCATCTCTGCTGCTCGGAGTCGCGATCCACTTCTTCAGGATGCTGTTGGCA 981
 Db 282 CCGTGGCGGCATCTCTGCTGCTCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 223
 QY 982 TTCCTTCGCTGCTTCTTACAGCCCTCTGCTGCTTCAAGAGCGCGGAGTAAAGACGACA 1041
 Db 222 TTGACTCCGCTGCTGATGTACAGCCCGCGCTTTCAGAGCGCGGCGATCACCGATGACG 163
 QY 1042 AACACTTCTTGGGACCACTTCGCGGTTCGCGGTTCACCAAGAGGCTTTCATCTGTTGG 1101
 Db 162 ACCAGCTTCTTGGCCACCCTTCGCGGTTCGCGGTTCACCAAGAGGCTTTCATCTGTTGG 103

QY	1102	CGACTTCTTCATCAGCGCGTCCGGCGCGCGCTGTGCTGGGACGACACGGGCGGGA	1161
Db	102	CCGCGGTCCTCCCTGGACCGCGTCCGGCAGGCGCCGCTGATGCTACAGCTCCGGCGGCA	43
QY	1162	TAATCTCTCTCCCTCATCGGCTCGCGCGCTCGCGCGCGGCTCACCGTCG	1203
Db	42	TGGCTGTCTCGCTCGTCCGCTCGCCACAGGCTCACCGTCG	1
RESULT 11			
CB620332			
LOCUS			
DEFINITION		671 bp mRNA linear EST 08-APR-2003	
		OSIIIEa05H07.f OSIIIEa Oryza sativa (indica cultivar-group) cDNA	
ACCESSION		clone OSIIIEa05H07 5', mRNA sequence.	
VERSION		CB620332	
KEYWORDS		CB620332.1 GI:29615320	
SOURCE		EST.	
ORGANISM		Oryza sativa (indica cultivar-group)	
		Oryza sativa (indica cultivar-group)	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
		Ehrhartoideae; Oryzeae; Oryza.	
REFERENCE		1 (bases 1 to 671)	
AUTHORS		Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,	
		Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.	
TITLE		Large-scale identification of ESTs involved in the interaction	
		between rice and Magnaporthe grisea	
JOURNAL		Unpublished (2003)	
COMMENT		Contact: Rod Wing	
		Arizona Genomics Institute	
		University of Arizona	
		Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ	
		85721-0088, USA	
		Tel: 520 626 3967	
		Fax: 520 621 9288	
		Email: http://genome.arizona.edu	
		PCR Primers	
		FORWARD: gta aaa cga cgg cca gtg	
		BACKWARD: gga aac agc tat gac cat g	
		Plate: 05 row: H column: 07	
		Seq primer: gta aaa cga cgg cca gtg.	
FEATURES			
source		Location/Qualifiers	
		1..671	
		/organism="Oryza sativa (indica cultivar-group)"	
		/mol_type="mRNA"	
		/cultivar="IR36"	
		/db_xref="taxon:39946"	
		/clone="OSIIIEa05H07"	
		/tissue_type="Leaf"	
		/dev_stage="3 week"	
		/lab_host="DH10B"	
		/clone_lib="OSIIIEa"	
		/note="vector: phluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Lesion Mimic SPL 11"	
ORIGIN			
Query Match		32.5%; Score 656.4; DB 14; Length 671;	
Best Local Similarity		98.8%; Pred. No. 1.7e-51;	
Matches 660; Conservative		1; Mismatches 7; Indels 0; Gaps 0;	
QY	33	CACCTCGACCGCCACTACTGTACACGGCCCGCAGAGCGAGCGCTCTCTCTGTGACACACCG	92
Db	4	CTCTGACCCCCCTTAAGTGTACAGCGCCCGCAGAGCGCTCTCTCTGTGACACACCG	63
QY	93	GAGATGCTTCGCGCGCTCGCGAGCGCCGTCCGCGGAGAAAGAGGGCAACGTCGG	152
Db	64	GAGATGCTTCGCGCGCGCTCGCGAGCGCGCTCGCGCTTAAGAGAAAGGGCAACGTCGG	123
QY	153	TTGCGCTTCGCGTCGCGCATCCTCGCTCCATGACCTCCATCTCTCTCGCTAGCATATC	212
Db	124	TTGCGCTTCGCGTCGCGCATCCTCGCTCCATGACCTCCATCTCTCTCGCTAGCATATC	183
QY	213	GGGGTGATGAGCGGGCGCTCGCTGTATCATCAAGAAGGACTTCAACATCAGTCAGCGGAAG	272

Best Local Similarity 99.4%; Pred. No. 2.8e-51; Matches 698; Conservative 0; Mismatches 0; Indels 4; Gaps 4;	
QY 1245	CTAGCATCGCTCCACCTCGCTACGCTCGCTTCTTCTCCATCGCGCTTGGCCCATC 1304
DB 1	CTAAGCATCGCTCCACCTCGCTACGCTCGCTTCTTCTCCATCGCGCTTGGCCCATC 60
QY 1305	ACGTGGGTACAGCTCGAGATCTTCCCGTCCAGGTGCGCGCTGGGTGCTCGCTC 1364
DB 61	ACGTGGGTACAGCTCGAGATCTTCCCGTCCAGGTGCGCGCTGGGTGCTCGCTC 120
QY 1365	GGCTGGCGGCAACCGGTCACAGCGGCTATCTCCATGACCTTCTGCTGCTGCC 1424
DB 121	GGCTGGCGGCAACCGGTCACAGCGGCTATCTCCATGACCTTCTGCTGCTGCC 180
QY 1425	AAGGCATCACATCGCGGCGAGCTTCTCTCTACTCCGATCGCGCTCGCTCG 1484
DB 191	AAGGCATCACATCGCGGCGAGCTTCTCTCTACTCCGATCGCGCTCGCTCG 240
QY 1485	GTGTTCTTCTACCTACTCTCCGAGACCGCGCGCGAGCGCTGGAGGATGAGCAAG 1544
DB 241	GTGTTCTTCTACCTACTCTCCGAGACCGCGCGCGAGCGCTGGAGGATGAGCAAG 300
QY 1545	CTGTTGCGGACACGGCGCGCGCC -TCGGAATCAGACGAGCCAGCAGAGAGAGAA 1603
DB 301	CTGTTGCGGACACGGCGCGCGCTTCGGAATCAGACGAGCCAGCAGAGAGAGAA 360
QY 1604	GGTGAATGCGCGCATCTAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1663
DB 361	GGTGAATGCGCGCATCTAATCAATCAATCAATCAATCAATCAATCAATCAATCA 420
QY 1664	TTTCTTCAAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1723
DB 421	TTTCTTCAAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 1724	GATTCGCTGATCCGCGCTTGTGGAGAGCGAGCGCGCGCGCGCGCGCGCGCGCG 1783
DB 481	GATTCGCTGATCCGCGCTTGTGGAGAGCGAGCGCGCGCGCGCGCGCGCGCGCG 540
QY 1784	CTGAGATCTTAAATCAATCTCAATGATGATGATGATGATGATGATGATGATGAT 1842
DB 541	CTGAGATCTTAAATCAATCTCAATGATGATGATGATGATGATGATGATGATGAT 600
QY 1843	CGTAAAGTGTACTATGCGATGAAGAAATACCATGTAGCAAGCTGAGGTGTGTG 1902
DB 601	CGTAAAGTGTACTATGCGATGAAGAAATACCATGTAGCAAGCTGAGGTGTGTG 659
QY 1903	TAGCTACTAGAGTGTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1944
DB 660	TAG-TACTAGAGTGTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
RESULT 13	
CB620395/c	
LOCUS	
DEFINITION	
OSiIEa05J08.1 r OSiIEa Oryza sativa (indica cultivar-group) cDNA	
clone OSiIEa05J08 3', mRNA sequence.	
ACCESSION	
CB620395	
VERSION	
CB620395.1	
KEYWORDS	
EST.	
SOURCE	
Oryza sativa (indica cultivar-group)	
Oryza sativa (indica cultivar-group)	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
Ehrhartoideae; Oryzaceae; Oryza.	
1 (bases 1 to 658)	
REFERENCE	
AUTHORS	
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,	
Kudrna,B., Dean,R., Soderlund,C., Wing,R. and Wang,G.	
TITLE	
Large-scale identification of ESTs involved in the interaction	
between rice and Magnaporthe grisea	
JOURNAL	
Unpublished (2003)	
COMMENT	
Contact: Rod Wing	
Arizona Genomics Institute	

University of Arizona Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA Tel: 520 626 3967 Fax: 520 621 9288 Email: http://genome.arizona.edu PCR Primers FORWARD: gta aaa cga cgg cca gtc g BACKWARD: gga aac agc tat gac cat g Plate: 05 row: J column: 08 Seq primer: gga aac agc tat gac cat g.	
FEATURES	
source	
1..858	
/organism="Oryza sativa (indica cultivar-group)"	
/mol_type="mRNA"	
/cultivar="IR36"	
/db_xref="taxon:39946"	
/clone="OSiIEa05J08"	
/tissue_type="Leaf"	
/dev_stage="3 week"	
/lab_host="DH10B"	
/clone_lib="OSiIEa"	
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Lesion Mimic SPL 11"	
ORIGIN	
Query Match 32.2%; Score 649; DB 14; Length 658;	
Best Local Similarity 99.1%; Pred. No. 8.3e-51;	
Matches 652; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY 1285	CCATCGCCCTTGGCCCATCACGTGGGTGTACAGTTCGGAGATCTTCCGCTCCAGGTGC 1344
DB 658	CCATCGCCCTTGGCCCATCACGTGGGTGTACAGTTCGGAGATCTTCCGCTCCAGGTGC 599
QY 1345	GGCGCTGGGCTGCTGCTCGCGCTCGCGCCAAACGGGTACACAGCGGCTCATCTCCA 1404
DB 598	GGCGCTGGGCTGCTGCTCGCGCTCGCGCCAAACGGGTACACAGCGGCTCATCTCCA 539
QY 1405	TGACCTTCTGCTGCTGCTCAAGGCATCACCATCGGGGAGCTTCTTCTCTACTCCG 1464
DB 538	TGACCTTCTGCTGCTGCTCAAGGCATCACCATCGGGGAGCTTCTTCTCTACTCCG 479
QY 1465	GCATCCCGCGCTCGCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCCGCGCGA 1524
DB 478	GCATCCCGCGCTCGCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCCGCGCGA 419
QY 1525	CGCTGGAGGATGAGCAAGCTGTTCCGGGACACGGCGCCCTCGGAATCAGACGAGC 1584
DB 418	CGCTGGAGGATGAGCAAGCTGTTCCGGGACACGGCGCCCTCGGAATCAGACGAGC 359
QY 1585	CAGCCAAGGAGAGAAAGGTGGAAATGGCCCACTAACTGATCAAACTAACCGCAA 1644
DB 358	CAGCCAAGGAGAGAAAGGTGGAAATGGCCCACTAACTGATCAAACTAACCGCAA 299
QY 1645	ATCACCATACTAAGGTTTCTTGCATAAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704
DB 298	ATCACCATACTAAGGTTTCTTGCATAAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 239
QY 1705	AGTAGCAGCAAGCTGGGAAGATTCTGATCCGGGTTGCTGGAGAGCGACCGCGCGGA 1764
DB 238	AGTAGCAGCAAGCTGGGAAGATTCTGATCCGGGTTGCTGGAGAGCGACCGCGCGGA 179
QY 1765	CGACAAAGCTGAGCTCCAGCTCGAGCTTCTTAAATCATCTTCAAGTACATGATTTTA 1824
DB 178	CGACAAAGCTGAGCTCCAGCTCGAGCTTCTTAAATCATCTTCAAGTACATGATTTTA 119
QY 1825	TTTGTCTTCTGCTTGTCCGTAAAGTTGTACTATGATGATGATGATGATGATGATG 1884
DB 118	TTTGTCTTCTGCTTGTCCGTAAAGTTGTACTATGATGATGATGATGATGATGATG 59
QY 1885	CAAGGCTGAGGTGTGTGTAGCTACTAGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1942
DB 58	CAAGGTTGAGGTGTGTGTAGCTACTAGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1

```

RESULT 14
CG260290/c
LOCUS CG260290 929 bp DNA linear GSS 25-AUG-2003
DEFINITION O3DF74TV ZM 0.7 1.5 KB Zea mays genomic clone ZM5BMa0780N04,
genomic survey sequence.
ACCESSION CG260290
VERSION CG260290.1 GI:34169668
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 929)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
CONSORTIUM for Maize Genomics
TITLE Unpublished (2002)
JOURNAL
COMMENT Other GSSs: CG3DF74TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
1..929
/mol_type="genomic DNA"
/organism="Zea mays"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMa0780N04"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 32.2%; Score 648.6; DB 29; Length 929;
Best Local Similarity 81.6%; Pred. No. 7.7e-51;
Matches 750; Conservative 0; Mismatches 169; Indels 0; Gaps 0;
Qy 635 CATCATCTCGGCATCGCGCGCGCGCTCGCTGCTCTCGCTCATGCTGCTCGGCAT 694
Db 929 CGTATGCTCGGCATCGCGCGCGCGCGCTCGCTGCTCTCGCTCATGCTGCTCGGCAT 870
Qy 695 GCCCGAGTCGCGCGCGCTGCTGCTCATGAAGGACGCTCGCGACGCCAAGGTGCT 754
Db 869 GCCCGAGTCGCGCTCGGTGCTGTGTCATGAAGGCGCGCTCGCGACGCCAGGCGCTGCT 810
Qy 755 GGAGAAGACCTCCACACGCGGAGGAGCGCGCGCGCGCTCGCGACGCCAAGGCGCG 814
Db 809 GGAGAAGACCTCCGCGCTCGCGGAGGAGCGCGCGCGCTCGCGACGCCAAGGTCG 750
Qy 815 CGCCGGATCCCTCAGGAGCTCGACGCGCGCGCTGTCACCGCTCCCGAAGAGAGGCGCG 874
Db 749 CGCGGGATCCGAAGGCGCTCGACGAGAGCTGTCGCGGTACCGAAGGAGCGCG 690
Qy 875 AAAAGAGAGCGGTGGAAGAGAGCTCATCTGTCGCCCGACCGCGCGCATCGCGCGCAT 934
Db 689 CGCGGAGATGAGGTGAGGAGAGCTCATCTTATCCCAACACCGCGCATCGCGCAT 630
Qy 935 CCGTGTGCGGGATCGGCATCCACTTCTCCACATGCGGTGGGCAATTCACCTCGGCT 994
Db 629 ACTGCTGTGCGCGCTGCGCTGCACTTCTTCCAGAGCGCTCTGGAGCGATTCGCGCT 570
Qy 995 CTTCTACAGCCCTCTCGGTTCGAAGAGCGCGCGGATTAAACGAACGACAAACACTTCTGGG 1054
Db 569 CCTGTACAGCCCGCGGTGTTCAAGAGCGCGGGGAATCACCGACGACAAACAGCTCTCTGGG 510

```

```

Qy 1055 CACCACTTGGCCGTTTCGGTGTCACCAAGAGGCTTTTCATCTCTGTGGGACTTTCTTCAT 1114
Db 509 CGTCACTTGGCGGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGT 450
Qy 1115 CGACCGCGTTCGGCGCGCGCGCTTCTGCGGACGACGCGCGGCGGATAATCTCTCCCT 1174
Db 449 GGACCGCGCGCGCGCTTCTGCTGACGACGCGCGGTGGAATGATGTTCTCGCT 390
Qy 1175 CATCGGCTTCGGCGCGCGGCTCACCGTGTGGCGAGCACCCCGACCGCAAGATACCTTG 1234
Db 389 CGTCTGCTTCGGGACGGGCTCACCGTGTGGGACCCACCGCGACGCAAGATCCCGTC 330
Qy 1235 GGCATTCGCGCTAACATCGCTCCACCTCGCTACGCTGCGCTTCTTCTCCATCGGCT 1294
Db 329 GCGGTCGCGCTGTGATCGCTCCACCTCGCTGCGCTTCTTCTCCATCGGCT 270
Qy 1295 TGGCCCATCATCGTGGGTGTACAGCTCGGAGATCTCCCGTCCAGGTGGCGCGCTGG 1354
Db 269 CGGCGCGCTCACGGGTGTGTACAACTCGGAGATATCCCGCTGAGGTGGCGCGCTGG 210
Qy 1355 CTGCTCGCTCGCGCTCGCGGCAACGGGTACACGCGGCTCATCTCCATGACCTTCT 1414
Db 209 ATTGCGGTTCGGCGTGGCGTGCACCGGTTACACGCGCGTTATCTCCATGACCTTCT 150
Qy 1415 GTGCTGTCCAAAGGCATCACCATCGCGGCGAGCTTCTTCTTACTCGGATCGCGCG 1474
Db 149 GTCCCTGTCCAAAGGTATCACCATCGCGGCGAGCTTCTTCTTACTCGGATCGCGCG 90
Qy 1475 GCTCGCTGGGTGTTCTTCTACACCTACTCCCGGAGACCGCGCGCGGACGCTGAGGA 1534
Db 89 GGTCTGGGTGGTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 30
Qy 1535 GATGAGCAAGCTGTCGCGC 1553
Db 29 GATGGCAAGCTGTCGCGC 11

RESULT 15
LOCUS CG064683/c
DEFINITION PUJCN07TD ZM_0.6_1.0_KB Zea mays genomic clone ZM5BMa0639B13,
genomic survey sequence.
ACCESSION CG064683
VERSION CG064683.1 GI:33936863
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 972)
AUTHORS Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PUCW07TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
1..972
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMa0639B13"

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:24:42 ; Search time 59 Seconds
(without alignments)
2456.727 Million cell updates/sec

Title: US-10-051-902a-20

Perfect score: 2559

Sequence: 1 MASDELAKAVEPRKGNVY.....AEEAEDAAAEKVVPELPSK 513

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2559	100.0	513	5	AAU97210 Corn Beta
2	2559	100.0	513	6	ABU08335 Corn suga
3	1906.5	74.5	529	5	AAU97214 Wheat sug
4	1906.5	74.5	529	6	ABU08339 Wheat sug
5	1872.5	73.2	510	5	AAU97211 Rice Beta
6	1872.5	73.2	510	6	ABU08336 Rice suga
7	1764.5	69.0	539	5	AAU97213 Wheat sug
8	1764.5	69.0	539	6	ABU08338 Wheat sug
9	1542	60.3	523	5	AAU97212 Soybean B
10	1542	60.3	523	6	ABU08337 Soybean S
11	1420	55.5	513	5	ABU08361 Amino aci
12	1316	51.4	491	3	AAU97207 Arabidops
13	1316	51.4	508	3	AAU97207 Arabidops
14	1245	48.7	466	3	AAU97207 Arabidops
15	772.5	30.2	333	6	ADA48320 Rice prot
16	678	26.5	487	6	ABG73334 Consensus
17	664.5	26.0	488	6	ABP98504 PFAM cons
18	664.5	26.0	488	6	AAE33506 Human sug
19	664.5	26.0	488	7	ADD22918 Human sug
20	646.5	25.3	629	4	AAE66935 Human GLU
21	646.5	25.3	629	4	AAE66940 Human GLU
22	637.5	24.9	648	5	ABG61548 Human tra
23	624.5	24.4	618	4	AAE66936 Rat GLUTX
24	611	23.9	446	6	ABU43281 Protein e
25	602	23.5	478	3	AAU15416 Arabidops

26	602	23.5	493	3	AAU15415 Arabidops
27	602	23.5	546	3	AAU15414 Arabidops
28	599	23.4	450	6	ABU33853 Protein e
29	597	23.3	582	3	AAU29528 Arabidops
30	595	23.3	555	3	AAU29529 Arabidops
31	593.5	23.2	480	4	AAU00107 Sugar tra
32	586.5	22.9	472	5	ABP52164 E. coli a
33	586.5	22.9	472	5	ABP52164 E. coli a
34	585	22.9	491	4	AAU89949 C glutami
35	583	22.8	491	4	AAU12594 Brevibact
36	576	22.5	167	5	AAU97209 Portion o
37	576	22.5	167	6	ABU08334 Corn suga
38	574.5	22.5	551	2	AAU50799 Spinach g
39	568	22.2	502	6	ABU36940 Protein e
40	565	22.1	584	5	ABU05603 Yeast ino
41	551.5	21.6	489	4	ABU63511 Drosophil
42	551.5	21.6	517	2	AAU49633 Wheat hex
43	548	21.4	517	5	ABP66239 Bifidobac
44	547	21.4	464	5	ABP52163 E. coli g
45	547	21.4	464	6	ABU14994 Protein e

ALIGNMENTS

RESULT 1
AAU97210
ID AAU97210 standard; protein; 513 AA.

XX AAU97210;

XX AC AAU97210;

XX DT 27-AUG-2002 (first entry)

XX DE Corn Beta vulgaris-like sugar transport protein.

XX KW Corn; Beta vulgaris-like sugar transport protein; carbohydrate transport;

XX KW grain filling; annual field crop; plant.

XX OS Zea mays.

XX PN US6383776-B1.

XX PD 07-MAY-2002.

XX PF 14-APR-1999; 99US-00291922.

XX PR 24-APR-1998; 98US-0083044P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

XX DR WPI; 2002-453364/48.

XX N-PSDB; ABK51971.

XX PT New nucleic acid encoding plant sugar-transport proteins, useful for

XX preparing transgenic plants with altered carbohydrate distribution.

XX PS Example 4; Fig 2; 54pp; English.

XX The present invention relates to the isolation of plant polynucleotide

XX sequences encoding an Arabidopsis thaliana-like sugar transport protein

XX or Beta vulgaris-like sugar transport protein. The polynucleotide

XX sequences are useful for altering the level of sugar transport proteins

XX in plants, i.e. for control of carbohydrate transport and distribution in

XX plant cells, e.g. during grain filling of annual field crops (e.g. corn,

XX rice, soybeans, and wheat), and, for studying carbohydrate flows and

XX sugar transport. The polynucleotide sequences can also be used to isolate

XX cDNA sequences and genes that encode homologues of the new proteins. The

XX present sequence represents a corn Beta vulgaris-like sugar transport

XX protein

XX Sequence 513 AA;

Apply. protein

```

Query Match      100.0%; Score 2559; DB 5; Length 513;
Best Local Similarity 100.0%; Pred. No. 3.1e-249;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASDELAKAVEPRKGNVYKASICAILASMASVILGYDGVMSGAAMYIKKDLNITDVL 60
DB 1 MASDELAKAVEPRKGNVYKASICAILASMASVILGYDGVMSGAAMYIKKDLNITDVL 60

QY 61 EILIGILSYLSFGSPAGARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120
DB 61 EILIGILSYLSFGSPAGARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120

QY 121 AGVGVGYGGMIAPVYTAISPAASRGFLTTPPEVFINIGILLGYLSNFAFARLPHLGWR 180
DB 121 AGVGVGYGGMIAPVYTAISPAASRGFLTTPPEVFINIGILLGYLSNFAFARLPHLGWR 180

QY 181 VMLAIGAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAERLADIKAA 240
DB 181 VMLAIGAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAERLADIKAA 240

QY 241 AGIPKGLDGDVTVPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOASGSDSV 300
DB 241 AGIPKGLDGDVTVPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOASGSDSV 300

QY 301 VOYSARLFKSAGITDDNKLGLVTCAGVTKTFFILVATFLDRAGRRLPLLISITGMIVS 360
DB 301 VOYSARLFKSAGITDDNKLGLVTCAGVTKTFFILVATFLDRAGRRLPLLISITGMIVS 360

QY 361 LICLSGLTVAGHPDTKVAVAVLCIASTLSYIAFFSIGLPTGVYTSIFFPQVRAL 420
DB 361 LICLSGLTVAGHPDTKVAVAVLCIASTLSYIAFFSIGLPTGVYTSIFFPQVRAL 420

QY 421 GFAVGASNRVTSAVISMTFLSKAITIGGSFFLYSGIAAVAVWFFFTCLPETRGRITL 480
DB 421 GFAVGASNRVTSAVISMTFLSKAITIGGSFFLYSGIAAVAVWFFFTCLPETRGRITL 480

QY 481 EMGKLCFMPDTCMAEEADAEEAAKEKVELPSSK 513
DB 481 EMGKLCFMPDTCMAEEADAEEAAKEKVELPSSK 513

RESULT 2
ABU08335 standard; protein; 513 AA.
XX AC ABU08335;
XX DT 29-MAY-2003 (first entry)
XX DE Corn sugar transport protein #3.
XX KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
XX KW plant sugar transport protein; carbohydrate transport; soybean;
XX KW carbohydrate distribution; plant.
XX OS Zea mays.
XX PN US2002178468-A1.
XX PD 28-NOV-2002.
XX PF 17-JAN-2002; 2002US-00051902.
XX PR 24-APR-1998; 98US-0083044P.
XX PR 14-APR-1999; 99US-00291922.
XX PA (ALLE/) ALLEN S M.
XX PA (HITZ/) HITZ W D.
XX PA (KINN/) KINNEY A J.
XX PA (TING/) TINGEY S V.
XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

```

```

XX MPI; 2003-340957/32.
DR N-PSDB; ABX93207.
XX PT Novel plant sugar transport proteins and nucleic acid encoding the
XX PT protein useful for producing transgenic plants having altered levels of
XX PT sugar transport protein.
XX PS Claim 10; Fig 2; 56pp; English.
XX The present invention relates to the isolation of Arabidopsis thaliana-
XX CC like or Beta vulgaris-like sugar transport proteins, and the
XX CC polynucleotide sequences encoding them. The plant sugar transport
XX CC proteins of the invention have been isolated from corn, rice, soybean,
XX CC and wheat. The polypeptides of the invention may be used for altering the
XX CC level of expression of a sugar transport protein in a host cell, by
XX CC transforming a host cell with a chimeric construct encoding all, or a
XX CC portion of the sugar transport protein, in sense or antisense
XX CC orientation. Particularly, the polypeptides may provide a means to
XX CC control carbohydrate transport and distribution in plants. ABU08334-
XX CC ABU08339 represent Beta vulgaris-like sugar transport proteins
XX SQ Sequence 513 AA;

Query Match      100.0%; Score 2559; DB 6; Length 513;
Best Local Similarity 100.0%; Pred. No. 3.1e-249;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASDELAKAVEPRKGNVYKASICAILASMASVILGYDGVMSGAAMYIKKDLNITDVL 60
DB 1 MASDELAKAVEPRKGNVYKASICAILASMASVILGYDGVMSGAAMYIKKDLNITDVL 60

QY 61 EILIGILSYLSFGSPAGARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120
DB 61 EILIGILSYLSFGSPAGARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120

QY 121 AGVGVGYGGMIAPVYTAISPAASRGFLTTPPEVFINIGILLGYLSNFAFARLPHLGWR 180
DB 121 AGVGVGYGGMIAPVYTAISPAASRGFLTTPPEVFINIGILLGYLSNFAFARLPHLGWR 180

QY 181 VMLAIGAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAERLADIKAA 240
DB 181 VMLAIGAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAERLADIKAA 240

QY 241 AGIPKGLDGDVTVPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOASGSDSV 300
DB 241 AGIPKGLDGDVTVPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOASGSDSV 300

QY 301 VOYSARLFKSAGITDDNKLGLVTCAGVTKTFFILVATFLDRAGRRLPLLISITGMIVS 360
DB 301 VOYSARLFKSAGITDDNKLGLVTCAGVTKTFFILVATFLDRAGRRLPLLISITGMIVS 360

QY 361 LICLSGLTVAGHPDTKVAVAVLCIASTLSYIAFFSIGLPTGVYTSIFFPQVRAL 420
DB 361 LICLSGLTVAGHPDTKVAVAVLCIASTLSYIAFFSIGLPTGVYTSIFFPQVRAL 420

QY 421 GFAVGASNRVTSAVISMTFLSKAITIGGSFFLYSGIAAVAVWFFFTCLPETRGRITL 480
DB 421 GFAVGASNRVTSAVISMTFLSKAITIGGSFFLYSGIAAVAVWFFFTCLPETRGRITL 480

QY 481 EMGKLCFMPDTCMAEEADAEEAAKEKVELPSSK 513
DB 481 EMGKLCFMPDTCMAEEADAEEAAKEKVELPSSK 513

RESULT 3
AAU97214 standard; protein; 529 AA.
XX ID AAU97214
XX AC AAU97214;
XX DT 27-AUG-2002 (first entry)
XX PI

```

DE Wheat sugar transport protein encoded by wlm1_pk0012.h1.
XX
KW Wheat; Beta vulgaris-like sugar transport protein;
XX carbohydrate transport; grain filling; annual field crop; plant.
XX
OS Triticum aestivum.
XX
XX US6383776-B1.
XX
PD 07-MAY-2002.
XX
XX 14-APR-1999; 99US-00291922.
XX
XX 24-APR-1998; 98US-0083044P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2002-453364/48.
XX
XX N-PSDB; ABK51975.
XX
XX New nucleic acid encoding plant sugar-transport proteins, useful for
XX preparing transgenic plants with altered carbohydrate distribution.
XX
XX Example 4; Col 77-80; 54pp; English.
XX
XX The present invention relates to the isolation of plant polynucleotide
XX sequences encoding an Arabidopsis thaliana-like sugar transport protein
XX or Beta vulgaris-like sugar transport protein. The polynucleotide
XX sequences are useful for altering the level of sugar transport proteins
XX in plants, i.e. for control of carbohydrate transport and distribution in
XX plant cells, e.g. during grain filling of annual field crops (e.g. corn,
XX rice, soybeans, and wheat), and, for studying carbohydrate flows and
XX sugar transport. The polynucleotide sequences can also be used to isolate
XX cDNA sequences and genes that encode homologues of the new proteins. The
XX present sequence represents a wheat Beta vulgaris-like sugar transport
XX protein
XX
SQ Sequence 529 AA;

Query Match 74.5%; Score 1906.5; DB 5; Length 529;
Best Local Similarity 73.6%; Pred. No. 2.7e-183;
Matches 373; Conservative 58; Mismatches 71; Indels 5; Gaps 3;
QY 1 MASDELAK--AVEPRKGNVYKASICAILASMASVILGYDYGMSGAAMYIKDLNITDV 58
DB 19 MASAAPEPGAVHPRNKGNFKYFTCALCASMATIVLGYDVGVMGASLYIKRDLQITDV 78
QY 59 QLEILIGILSLYLFSGFAGARTSDRIGRLTVVFAAVIFVGSLLMGFAVNYGMLMAGR 118
DB 79 QLEIMMGILSYALIGSLGARTSDWYGRVTVVFAAAIFNNGSLMGFAVNYAMLMVGR 138
QY 119 FVAGVGVGGMIAPVVTAETSPAASRGFLTTPPEVFINTGILLGYLSNFAFARLPHLG 178
DB 139 FVTGIGVYALMVAVVTPPEVSPASARGFLTSTFEVFINGVILLGYSNVAFARLPHL 198
QY 179 WRVMLATGAVPSGLLALLVFCMPSPRWLVKGLADARAVLEKTSATPEAAERLADIK 238
DB 199 WRVMLGICAVPSALLVLMVFCMPSPRWLVKGLADARAVLAKTSDTPEAEVERLDQIK 258
QY 239 AAGIGPKGLDGVVTVFGKGGGELQVWKLLILSPTPAVRILLISAVGHFFQOASGSD 298
DB 259 AAGIGPRELDGVVTVMP-KTGGQEKQVWKELIPSPTPAMRRILLALGHHFFQOATGSD 317
QY 299 SWOYSARLFKSAGITDNDKLLGVTCVAGVTKTFILVATFLDLDRGRRLILLISTGMI 358
DB 318 SVLVYSPRVFQSAGITGDNHLLGATCANGVYKTLFILVATFQLDVRGRRPLLTSTAGML 377
QY 359 VSLICLSGLITVAGHHFDPKVAWVALCIATSLSYIAFFSIGLPITGVTSIFPLQVR 418
DB 378 ACLIGLGTGLTVVGRHPDAKVPWAIGLCIVSYLAVVFFSIGLPITSVYTFVPLVRV 437

QY 419 ALGPAVGASNEVTSNAVISMTFLSLSKAITIGSFYLSGIAAVAVWVFFTCLPETRGR 478
DB 438 ALGFALGTSCNRVTSAAVMSFSLSKAITIGSFYLAGIAAGWIFFFTFIPETRGLP 497
QY 479 LEEEMKGLFGMPDTGMAREAEADAAAEK 505
DB 498 LEEIGKLFQMTDT--AVEAQDTATDK 522
RESULT 4
ABU08339
ID ABU08339 standard; protein; 529 AA.
XX
XX AC ABU08339;
XX
XX DT 29-MAY-2003 (first entry)
XX
XX DE Wheat sugar transport protein #5.
XX
XX KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
XX plant sugar transport protein; carbohydrate transport; soybean;
XX carbohydrate distribution; plant.
XX
XX OS Triticum aestivum.
XX
XX PN US2002178468-A1.
XX
XX PD 28-NOV-2002.
XX
XX PF 17-JAN-2002; 2002US-00051902.
XX
XX PR 24-APR-1998; 98US-0083044P.
XX
XX PR 14-APR-1999; 99US-00291922.
XX
XX PA (ALBE/) ALLEN S M.
XX (HITZ/) HITZ W D.
XX (KINN/) KINNEY A J.
XX (TING/) TINGEY S V.
XX
XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2003-340957/32.
XX N-PSDB; ABX93211.
XX
XX Novel plant sugar transport proteins and nucleic acid encoding the
XX protein useful for producing transgenic plants having altered levels of
XX sugar transport protein.
XX
XX Claim 10; Fig 2; 56pp; English.
XX
XX The present invention relates to the isolation of Arabidopsis thaliana-
XX like or Beta vulgaris-like sugar transport proteins, and the
XX polynucleotide sequences encoding them. The plant sugar transport
XX proteins of the invention have been isolated from corn, rice, soybean,
XX and wheat. The polypeptides of the invention may be used for altering the
XX level of expression of a sugar transport protein in a host cell, by a
XX transforming a host cell with a chimeric construct encoding all, or a
XX portion of the sugar transport protein, in sense or antisense
XX orientation. Particularly, the polypeptides may provide a means to
XX control carbohydrate transport and distribution in plants. ABU08334-
XX ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ Sequence 529 AA;
Query Match 74.5%; Score 1906.5; DB 6; Length 529;
Best Local Similarity 73.6%; Pred. No. 2.7e-183;
Matches 373; Conservative 58; Mismatches 71; Indels 5; Gaps 3;
QY 1 MASDELAK--AVEPRKGNVYKASICAILASMASVILGYDYGMSGAAMYIKDLNITDV 58
DB 19 MASAAPEPGAVHPRNKGNFKYFTCALCASMATIVLGYDVGVMGASLYIKRDLQITDV 78
QY 59 QLEILIGILSLYLFSGFAGARTSDRIGRLTVVFAAVIFVGSLLMGFAVNYGMLMAGR 118

```

Db 79 QLEINMGLSLVYALGSLGARTSDWGRRTVVVFAAAIFNNGSLLMGFAVNYAMLMVGR 138
Qy 119 FVAGVGVGGMIAPIVYTAETSPASARGFLTFTEVEFINIGILGYLSNFAFARLPHLG 178
Db 139 FVTGIGVYAIMVAPVYTPVSPASARGFLTFTEVEFINIGILGYLSNFAFARLPHLG 198
Qy 179 WRVMAIGAVPSGLLALLVFCMPSPRLVJLKGRLADARAVLEKTSATPBEAERLADIK 238
Db 199 WRVMLGIGAVPSALLVFCMPSPRLVJLKGRLADARAVLEKTSATPBEAERLADIK 258
Qy 239 AAAGIPKGLDGVTVPCKEGGELQVWKKLILSPVAVRRILLSSAVGLHFFQOAGSD 298
Db 259 AAAGIPRELDDGVVMP-KTKGQGEKQVWKELISPTPAMERILLALGHLHFFQOAGSD 317
Qy 299 SVQVYASRLFSAGITDNKLLGVTCVAVGVTKTFPILVATFLDRAGRRPLLLISTGMI 358
Db 318 SVVLYSPRVFSAGITDNHLLGATCANGVMKTLFILVATFLDRAGRRPLLLISTGML 377
Qy 359 VSLICGLSLTVAGHHPDTKVAVAVLCIASTLSVIAFFSGLGPIGCVTSEIFPLQVR 418
Db 378 ACLIGLGLTVAGHHPDTKVAVAVLCIASTLSVIAFFSGLGPIGCVTSEIFPLQVR 437
Qy 419 ALGFAVGAVSNRVTSAVISMFTLSLKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGR 478
Db 438 ALGFAVGAVSNRVTSAVISMFTLSLKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGR 497
Qy 479 LEEMKGLFGMPDTGMABEADAARKEK 505
Db 498 LEEIGKLFGMTDT--AVEAQDTATKDK 522

```

RESULT 5

```

AAU97211
ID AAU97211 standard; protein; 510 AA.
XX
AC AAU97211;
XX
DT 27-AUG-2002 (first entry)
XX
DE Rice Beta vulgaris-like sugar transport protein.
XX
KW Rice; Beta vulgaris-like sugar transport protein; carbohydrate transport;
XX grain filling; annual field crop; plant.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT Misc-difference 102
FT /label= Unknown
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2002-453364/48.
DR N-PSDB; ABK51972.
XX
XX New nucleic acid encoding plant sugar-transport proteins, useful for
XX preparing transgenic plants with altered carbohydrate distribution.
XX
XX Example 4; Fig 2; 54pp; English.
XX
XX The present invention relates to the isolation of plant polynucleotide
XX sequences encoding an Arabidopsis thaliana-like sugar transport protein

```

```

CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence represents a rice Beta vulgaris-like sugar transport
CC protein
XX
SQ Sequence 510 AA;

```

```

Query Match 73.2%; Score 1872.5; DB 5; Length 510;
Best Local Similarity 72.7%; Pred. No. 7, 1e-180;
Matches 372; Conservative 60; Mismatches 77; Indels 3; Gaps 2;
Qy 1 MASDELAKAVEPRKXGNVYASICAILASMASVILGYDIGNVSGAAMIKKDLNITDVL 60
Db 1 MASAALPEAVAPKXGNVRFACAILASMTSILLGYDIGNVSGASLYIKKDFNISDGKV 60
Qy 61 EILIGLSLTVAGHHPDTKVAVAVLCIASTLSVIAFFSGLGPIGCVTSEIFPLQVR 120
Db 61 EVLMGLNDYSLIGSFAAGRTSDWGRRTVVVFAAAIFNNGSLLMGFAVNYAMLMVGR 120
Qy 121 AGVGVGYGGMIAPIVYTAETSPASARGFLTFTEVEFINIGILGYLSNFAFARLPHLQWR 180
Db 121 AGIGVGYALMIAPIVYTAETSPASARGFLTFTEVEFINIGILGYLSNFAFARLPHLQWR 180
Qy 181 VMLAIGAVPSGLLALLVFCMPSPRLVJLKGRLADARAVLEKTSATPBEAERLADIKA 240
Db 181 IMLGGAAPSVLLALVLMGMPSPRLVJLKGRLADARAVLEKTSATPBEAERLADIKA 240
Qy 241 AGIPKGLDGVTVPCKEGGELQVWKKLILSPVAVRRILLSSAVGLHFFQOAGSDSV 300
Db 241 AGIPEELDGVTVTP-KRSGNEKRVWKELILSPVAVRRILLSSAVGLHFFQOAGSDSV 299
Qy 301 VOYSARLFKSAGITDNKLLGVTCVAVGVTKTFPILVATFLDRAGRRPLLLISTGMI 360
Db 300 VFYSPLVFKSPGLTKHFLGTTNPFVGTNKLIFLLATFFIDGVGRRLLLGSGTGII 359
Qy 361 LILCGLSLTVAGHHPDTKVAVAVLCIASTLSVIAFFSGLGPIGCVTSEIFPLQVR 420
Db 360 LILCGLSLTVAGHHPDTKVAVAVLCIASTLSVIAFFSGLGPIGCVTSEIFPLQVR 419
Qy 421 GFAVGAVSNRVTSAVISMFTLSLKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGR 480
Db 420 GCSLGVAAANRVTSGVISMFTLSLKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGR 479
Qy 481 EMCKLFGMPDTGMABEADAARKEKVELPSS 512
Db 480 EMCKLFG--DTAAASESDPEAKEKKVMAAT 509

```

RESULT 6

```

ABU08336
ID ABU08336 standard; protein; 510 AA.
XX
AC ABU08336;
XX
DT 29-MAY-2003 (first entry)
XX
DE Rice sugar transport protein #3.
XX
KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.
XX
OS Oryza sativa.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX

```